Journal of Safety Research 76 (2021) 127–134

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

A comparison of general aviation accidents involving airline pilots and instrument-rated private pilots

Douglas D. Boyd *, Mark Scharf, David Cross

Embry-Riddle Aeronautical University, 1 Aerospace Blvd, Daytona Beach, FL 32114, USA

A R T I C L E I N F O

Article history:
Received 23 April 2020
Received in revised form 21 July 2020
Accepted 19 November 2020
Available online 15 December 2020

Keywords:
General aviation
Light aircraft
Flying accidents
Airline pilots
Aviation safety

A B S T R A C T

Introduction: The extremely low accident rate for U.S air carriers relative to that of general aviation (~1 and ~60/million flight hours respectively) partly reflects advanced airman certification, more demanding recurrency training and stringent operational regulations. However, whether such skillset/training/regulations translate into improved safety for airline pilots operating in the general aviation environment is unknown and the aim of this study. *Methods:* Accidents (1998–2017) involving airline pilots and instrument-rated private pilots (PPL-IFR) operating non-revenue light aircraft were identified from the NTSB accident database. An online survey informed general aviation flight exposure for both pilot cohorts. Statistics used proportion testing and Mann-Whitney *U* tests. *Results:* In degraded visibility, 0 and 40% (χ^2 $p = 0.043$) of fatal accidents involving airline and PPL-IFR airmen were due to in-flight loss-of-control, respectively. For landing accidents, airline pilots were under-represented for mishaps related to airspeed mismanagement ($p = 0.036$) relative to PPL-IFR but showed a dis-proportionate count (2X) of ground loss-of-directional control accidents ($p = 0.009$) the latter likely reflecting a preference for tail-wheel aircraft. The proportion of FAA rule violation-related mishaps by airline pilots was >2X (7 vs. 3%) that for PPL-IFR airmen. Moreover, airline pilots showed a disproportionate (χ^2 $p = 0.021$) count of flights below legal minimum altitudes. Not performing an official preflight weather briefing or intentionally operating in instrument conditions without an IFR flight plan represented 43% of airline pilot accidents involving FAA rule infractions. *Conclusions:* These findings inform safety deficiencies for: (a) airline pilots, landing/ground operations in tail-wheel aircraft and lack of 14CFR 91 familiarization regulations regarding minimum operating altitudes and (b) PPL-IFR airmen in-flight loss-of-control and poor landing speed management. *Practical Applications:* For PPL-IFR airmen, training/recurrency should focus on unusual attitude recovery and managing approach speeds. Airline pilots should seek additional instructional time regarding landing tail-wheel aircraft and become familiar with 14CFR 91 rules covering minimum altitudes.

1. Introduction

Civil aviation can be arbitrarily divided into (a) revenue-based transportation comprised mainly of air carrier operations utilizing transport-category aircraft (>12,500 lbs.) and (b) general aviation employing light aircraft (≤12,500 lbs.) (Boyd, 2017). While air carrier operations have, over the last few decades, boasted a stellar safety record (Boyd, 2017), alas general aviation, despite a modest decrease in accident rate over most recent years, still shows a lackluster record with a >60 times higher accident (herein also referred to as mishaps) rate (Boyd, 2017; Li & Baker, 2007).

The discrepancy in safety between airline and general aviation operations probably reflects multiple factors. First is the advanced certification and recurrency training requirements for airline aircrews. Presently, to exercise flying privileges for an air carrier, pilots must be air-transport pilot (ATP) certificated (Electronic Code of Federal Regulation, 2020) whereas for general aviation operations the majority of general aviation airmen (Federal Aviation Administration, 2015) hold a private pilot (PPL) certificate. In this regard, greater precision in regard to both instrument flight (i.e., operating the aircraft by sole reference to flight instruments) (Federal Aviation Administration, 2019, 2018) and landing operations are demanded for the ATP certificate. Specifically for instrument flight, a one quarter versus a three quarter scale lateral deflection of the course deviation indicator is allowed for the ATP (Federal Aviation Administration, 2019) and private pilot instru-

* Corresponding author at: 1014 Sugar Mountain Court, Sugar Land, TX 77498, USA.

E-mail address: boydd8@erau.edu (D.D. Boyd).

https://doi.org/10.1016/j.jsr.2020.11.009
0022-4375/

ment ratings (Federal Aviation Administration, 2018) respectively. Similarly, landing operations have tighter tolerances for ATP certification (a 100 vs. 200 foot margin for spot landings). Transport-category aircraft spot landings (to mitigate against the possibility of a runway overrun) require precise energy-management (Federal Aviation Administration, 2016, 2008) due to greater landing distances required than a light aircraft. Recurrency training for air carrier pilots is also more frequent and demanding compared with general aviation (Electronic Code of Federal Regulation, 2015). Crews have to undertake such training every 6 (Captain) or 12 months (first officer) whereas a flight review for general aviation airmen operating light aircraft for non-revenue is only required once every 24 months (Electronic Code of Federal Regulation, 2018). Moreover, recurrency programs for airline pilots are more extensive typically consisting of a multi-day program (comprised of maneuvers, abnormal procedures, upset recoveries and line-oriented flight training (Electronic Code of Federal Regulation, 2015). In contrast a flight review for a PPL requires only 1 hour of flight and tasks are at the sole discretion of the instructor "as necessary for safe flight" (Electronic Code of Federal Regulation, 2018). A second reason for the superior safety of the air carriers is the more stringent regulations (14CFR 121) (Electronic Code of Federal Regulation, 2017) governing their operations (relative to the corresponding rules (14CFR 91) governing general aviation (Electronic Code of Federal Regulation, 2015) as well as the use of standard operating procedures (Administration, 2014) the latter absent from general aviation. For instance, while airport minimum visibility requirements apply to departing air carrier flights (14CFR 121.637), no such restrictions limit general aviation (14CFR 91) operations (Electronic Code of Federal Regulation, 2017). Third, although not mandatory, many U.S. carriers have adopted safety management systems (SMS) and threat and error management training per Federal Aviation Administration recommendations (Federal Aviation Administration, 2012, 2016). Lastly, aircraft employed for air carrier operations are certificated (14CFR 25) to a higher safety standard (Electronic Code of Federal Regulation, 2017) with a greater level of equipment redundancy than airplanes (14 CFR 23) (Electronic Code of Federal Regulation, 2012) used in general aviation.

Nevertheless, for the airline pilot operating light aircraft under 14CFR 91, certain aspects of air carrier operations could potentially offset the safety-promoting factors cited above. For example, automation, more prevalent for transport-category aircraft, has raised concern as to the erosion of manual flying skills with one research study demonstrating degraded Boeing 747 pilot performance when tasked with manual flying (Casner, Geven, Recker, & Schooler, 2014). In addition, the typical general aviation light aircraft requires more control inputs of the primary flight control surfaces for any given wind conditions than a much heavier transport-category airplane subjected to identical conditions. Lastly, virtually all transport-category aircraft employed by air-carriers require two person crews (14CFR 25 certification (Electronic Code of Federal Regulation, 2017) allowing for a prescriptive division of tasks for the pilot flying and pilot monitoring (14CFR 121.542-545 (Electronic Code of Federal Regulation, 2017)). In contrast, the vast majority of light aircraft are operationally approved for, and piloted, by a single crew member (Electronic Code of Federal Regulation, 2017) with an attendant increase in workload (Electronic Code of Federal Regulation, 2017).

Thus, whether more rigorous airman certification/recurrency training/stringent operational rules for airline pilots translates into improved safety in the general aviation environment or conversely, whether lesser automation coupled with lighter aircraft performance (more subject to winds) offsets such safety benefits has yet to be determined. Accordingly, we undertook a study to determine the level of safety of airline pilots flying non-revenue, light aircraft in operational areas where their professional training/experience/regulations, as described above, would be expected to impact. Specifically, the following question was posed: are airline pilots superior to their instrument-rated private pilot (PPL-IFR) counterparts as evidenced by a reduced proportion of: (a) accidents attributed to an in-flight loss-of-control in degraded visibility- an event (International Air Transport Association, 2018) previously cited on the NTSB "Most Wanted List" (National Transportation Safety Board, 2016); (b) landing accidents ascribed to deficient pilot technique; and (c) mishaps involving violation of the general aviation operations regulations (14CFR 91).

2. Materials and methods

2.1. Procedure

Accidents were identified from a retrospective search of the downloaded NTSB Microsoft Access database (2018 Oct release) (National Transportation Safety Board, 2015) involving: (a) airline pilots the latter defined as an ATP-certificated professional airman holding a Class 1 medical, a type rating in a transport-category aircraft (or employed by an air carrier) and 65 years or younger; and (b) as a control group, instrument-rated PPLs holding a Class 3 medical. It should be noted that the PPL population was deliberately restricted to those airmen concurrently holding an IFR rating (hereafter referred to as PPL-IFR pilots) to afford a comparison for airman performance in degraded visibility – instrument flying proficiency representing a core element of the ATP certificate (Federal Aviation Administration, 2019).

The database was queried for accidents occurring over the period spanning 1998–2017 involving piston engine-powered airplanes ($\leq$12,500 lbs.) in which flights were conducted under general operating flight rules (14CFR 91 (Electronic Code of Federal Regulation, 2015)) for personal missions. Accidents in Alaska were excluded from the query strategy. Data were exported to Excel and checked for duplicates (which were deleted). Accident causes were per the NTSB final report. Airline pilot type rating data was obtained from a variety of publicly available resources (Aviation DB, 2019; Federal Aviation Administration, 2019) and by the FAA Office of Accident Investigation and Prevention.

High-energy landings were defined as those for which the NTSB final report cited porpoising, multiple bounces or floating of the accident airplane (Boyd, 2019; Goode, O'Bryan, Yenni, Cannaday, & Mayo, 1976). Conversely, landings with inadequate airspeed (low-energy) were those cited as such or for which an aerodynamic stall occurred above the runway again per the NTSB final report (Boyd, 2019).

An anonymous online survey as to non-revenue, 14CFR 91 operations of light aircraft by PPL-IFR and airline pilots (approved by the Embry Riddle Aeronautical University Institutional Review Board) to inform flight times and ambient conditions was constructed in SurveyMonkey^{R} (www.surveymonkey.com) and pretested by four FAA Safety Team general aviation pilots as well as co-authors MS and DC. Responses from the airline and PPL-IFR pilot populations at large were collected over the period spanning Feb 14-April 05, 2020.

2.2. Statistical analysis

Proportion testing used contingency tables and a Pearson Chi-Square or Fisher's Exact (2-sided) tests to determine where there were statistical differences (Agresti, 2012; Field, 2009). The contribution of individual cells in proportion tests was determined using standardized residuals (Z-scores) in post-hoc testing. Differences in median values for non-normally distributed data (determined using a Shapiro Wilks test) were tested using a Mann-Whitney test. All statistical analyses were performed using SPSS (v24) software.

3. Results

3.1. Accident pilot population

In the retrospective analysis, a query of the NTSB Access database for general aviation accidents in the United States involving light aircraft occurring over the period spanning 1998–2017 returned 124 and 934 airline and PPL-IFR pilots with median ages of 49 and 54 years, respectively. These two airman cohorts had accrued a median total flight experience of 12,917 and 1,042 hours in all aircraft, respectively.

3.2. In-flight loss-of-control accidents in degraded visibility

We argued, that with the greater precision required for instrument flight per ATP certification (Federal Aviation Administration, 2019, 2018) and an increased exposure to degraded visibility concomitant with their professional occupation a reduced proportion of in-flight loss-of-control accidents in such visibility would be evident for airline pilots. Herein, degraded visibility was operationally defined as less than visual flight rules (i.e. cloud ceiling of ≤3,000 feet (AGL)) and/or ambient night lighting (Federal Aviation Administration, 2018).

While 27% of instrument-rated private pilot (PPL-IFR) accidents occurred in degraded visibility (Fig. 1), airline pilots showed a lower proportion (11%) of such mishaps, a difference which was statistically significant (χ^2 $p < 0.001$). Loss-of-control accidents often have a fatal outcome (National Transportation Safety Board, 2018) and indeed, this cause was previously cited on the NTSB "most wanted" list (National Transportation Safety Board, 2016). Perhaps not surprisingly, 40% of fatal accidents in degraded visibility involving PPL-IFR airmen were ascribed to this event (Fig. 2). In contrast, for airline pilots, there were no accidents in degraded visibility attributed to in-flight loss-of-control. Again, this difference in proportions between the two pilot groups was statistically significant (χ^2 $p = 0.043$).

To determine if this absence of in-flight loss-of-control accidents incurred by airline pilots was due to diminished general aviation flying in degraded visibility, the air carrier and PPL-IFR pilot populations at large were, in a prospective online survey, queried for their flight times and environmental conditions whilst operating light aircraft under 14CFR 91. Of 913 respondents, 295 airline and 618 PPL-IFR airmen completed the survey (Table 1). While indeed, the latter airmen showed an approximately 3 fold increase in annual IMC/night flight times compared with air carrier pilots, this difference unlikely accounts for the complete lack of in-flight loss-of-control accidents involving airline pilots operating in degraded visibility.

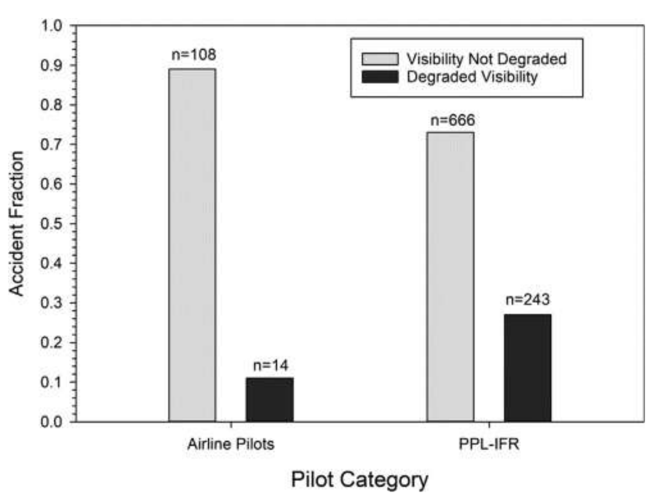


Fig. 1. Accidents in degraded visibility. The proportion of light aircraft accidents in degraded visibility is shown for each pilot cohort. Degraded visibility was operationally defined as a cloud ceiling of equal to or less than 3000 ft. AGL or ambient night lighting. Proportion differences were tested with a 2-sided Chi-Square test. PPL-IFR, instrument rated private pilots, n = accident count.

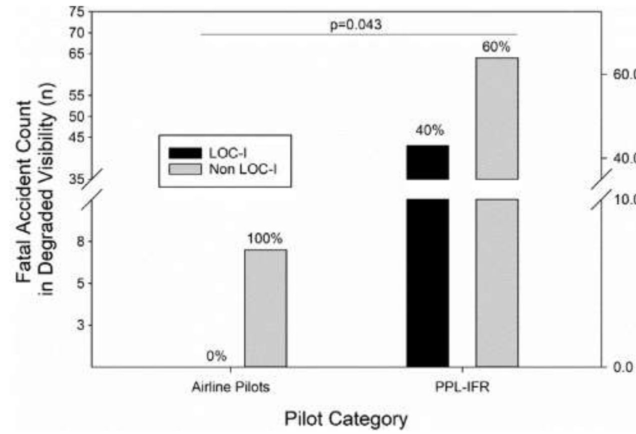


Fig. 2. Fatal in-flight loss-of-control accidents in degraded visibility. Fatal accident counts in degraded visibility (as described in Fig. 1) due to in-flight loss-of-control (LOC-I) or other causes (Non LOC-I) for the specified pilot category are shown. The difference in proportions of accidents was tested for with a Fisher's Exact Test.

3.3. Landing accidents

Landing a transport category aircraft requires a higher degree of precision than a comparable operation with a light aircraft due to the greater weight and physical dimensions. Specifically, a substantially higher weight (e.g., maximum landing weight of a Boeing 737–800 is 146,275 lbs. (Modern Airliners, 2020)) 57 fold higher than that of a Cessna 172S (2,550 lbs.) (Cessna, 2020) necessitates a faster landing speed that must be closely adhered to in order to avoid a runway overrun. Likewise, the greater lateral spacing of the main landing gear wheels also demands precision in directional control of a transport-category aircraft after touchdown. In contrast, operating a light aircraft at the majority of U.S. civil aviation airports (Federal Aviation Administration, 2017) with their relatively long and wide runways allow for deficiencies in the aforementioned skills with a reduced risk of a runway excursion. To determine if the airline pilot landing skillset transferred to the operations of light aircraft, landing accidents were compared for the two pilot cohorts.

Across all phases of flight operations, landing accidents were the most frequent for both airline and PPL-IFR pilots accounting for 39% ($n = 22$) and 27% ($n = 259$) of mishaps, respectively. Although the elevated proportion for air carrier airmen relative to the PPL-IFR cohort was not statistically different (χ^2 $p = 0.069$), nevertheless, it contravenes the notion that landing proficiency skills in transport-category aircraft transfers to light aircraft operations.

Landing accidents ascribed to deficiencies in pilot stick and rudder skills were then categorized as to cause. Airline pilots were superior to their PPL-IFR counterparts in energy management with zero landing mishaps ascribed to either excessive (High-energy Approach) or insufficient speed (Low-Energy Approach) (Fig. 3). On the other hand, approximately 26% of landing mishaps by

Table 1
Prospective survey of airline and PPL-IFR pilots.

		Airline Pilots	PPL-IFR	*P* value
Age	*n*	295	618	
	Median (h)	53	60	<0.001
	Q1 (h)	43	48	
	Q3 (h)	60	68	
Total Time Light Aircraft (h) Most Commonly Flown	*n*	293	615	
	Median (h)	500	730	0.011
	Q1 (h)	260	323	
	Q3 (h)	1350	1700	
Annual Light Aircraft Time (h) Airplane Most Frequently Flown	*n*	292	613	
	Median (h)	75	80	0.005
	Q1 (h)	44	50	
	Q3 (h)	100	120	
Last 90 Days Flight Time (Make-Model)	*n*	295	618	
	Median (h)	10	16	<0.001
	Q1 (h)	5	9	
	Q3 (h)	20	30	
Number of Flights in Light Aircraft Most Commonly Operated Last 90 Days	*n*	295	618	
	Median	8	10	0.069
	Q1	3	5	
	Q3	15	17	
Annual IMC Time (h) Light Aircraft Most Frequently Operated	*n*	295	618	
	Median (h)	2	7	<0.001
	Q1 (h)	0	2	
	Q3 (h)	10	15	
Annual Night Time (h) Light Aircraft Most Commonly Used	*n*	295	618	
	Median (h)	2	7	0.005
	Q1 (h)	0	2	
	Q3 (h)	10	15	
Landing Gear Type	Tail-Wheel (*n*)	90	32	$p < 0.001$
	%	31	5	
	Nose-Wheel (*n*)	205	586	
	%	69	95	

Results of an online survey conducted of the airline (Airline) and instrument-rated private (PPL-IFR) pilot population-at-large. Data were non-normally distributed per a Shapiro-Wilk test and accordingly differences in median values tested using a Mann-Whitney *U* Test. Proportion differences for landing gear type was tested using a Chi-Square test. h, hours; Q, quartile. IMC, instrument meteorological conditions.

PPL-IFR airmen were due to a high-energy approach (defined as any in which the aircraft porpoised, floated or bounced multiple times) a difference which was statistically significant (χ^2 $p = 0.036$).

Conversely, a higher (χ^2 $p = 0.009$) proportion of landing accidents which the NTSB binned into the "ground loss of directional control" category (0.92 vs. 0.53) was evident for airline pilots (Fig. 3). This was unexpected as managing this vector component is more critical for a transport category aircraft with its substantially wider main wheel base compared with that of a light aircraft. We considered the possibility that this surprising finding was related to the type of aircraft landing gear. Tail-wheel (conventional) and tricycle (nose) landing gear-equipped aircraft exhibit different handling characteristics and are well recognized as more challenging to maintain ground directional control particularly in a cross-wind (Kirby, 2020). Indeed, consistent with this argument, while over 70% of landing accidents involving air carrier pilots were incurred with tail-wheel airplanes, this proportion was substantially lower (<30%) for such mishaps involving PPL-IFR airmen (Fig. 4). This difference in accident aircraft landing gear type was statistically significant for the two pilot cohorts (χ^2 $p < 0.001$). Presumably, the over-representation of this type of landing accident for airline pilots reflects their preference for such-equipped aircraft for general aviation operations (Table 1). It is worth noting that none of the ground loss of directional control accidents involving airline pilots in tail-wheel equipped airplanes was due to an exceedance of the maximum demonstrated cross-wind component.

3.4. Violation of FAA regulations

Airline operations are under strict vigilance for infringement of the FAA regulations via a variety of mechanisms including flight quality assurance programs (Federal Aviation Administration, 2004) and audio recordings of the flight deck (Electronic Code of Federal Regulation, 2017). In contrast, little comparable oversight exists for general aviation. Moreover, airline pilots are well aware that infractions of the regulations leading to an incident or accident may culminate in the revocation of flying privileges and hence income. With these factors in mind, we hypothesized that a diminished fraction of FAA violation-related general aviation accidents would be evident for these airmen whilst operating light aircraft.

Contrary to expectations, the proportion of 14CFR 91 rules transgression-related mishaps by airline pilots, although low, was more than double (7 vs. 3% respectively) that for accidents involving PPL-IFR airmen. There was little evidence of a temporal trend in such accidents as 3 and 4 of mishaps involving an infraction of the FAA regulations occurred over the 1998–2007 and 2008–2017 periods respectively, The infractions of the FAA regulations were then sub-categorized (Table 2). Interestingly, there was a disproportionate (χ^2 $p = 0.021$) number of violations involving airline pilots in which the light aircraft was operated below the legal minimum altitude –accounting for 57% of FAA rule infractions. In contrast, this subcategory accounted for 17% of all PPL-IFR accidents in which the FAA regulations were breached. Interestingly, the second most common (constituting 43% of all FAA

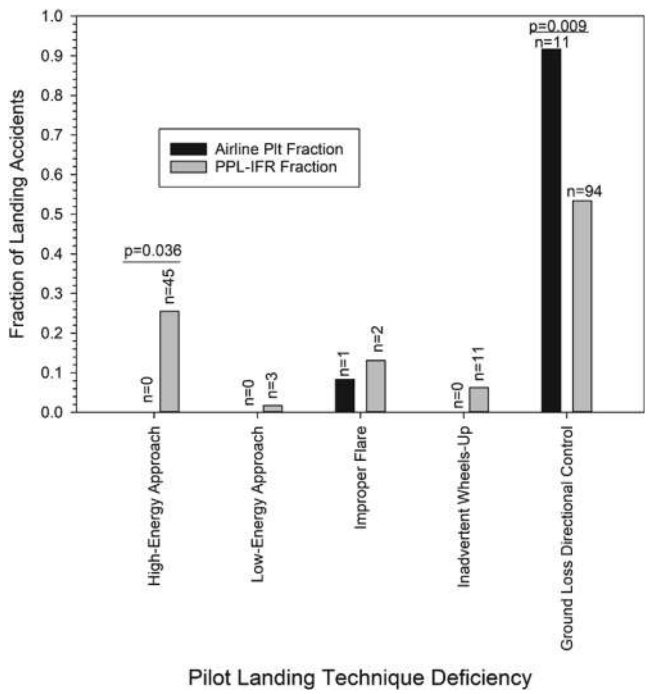


Fig. 3. Causes of landing accidents related to deficiency in pilot technique. The proportion of landing accidents for the specified causes related to pilot stick-and-rudder skills is illustrated as a function of pilot category. A high-energy landing accident was one in which the NTSB report cited porpoising, floating or multiple bounces of the aircraft. A low-energy landing mishap was one in which the aircraft stalled above the runway. The *p* values shown were derived from adjusted residuals of a Chi-Square test. *n*, count of accidents.

transgressions) violation of the FAA regulations for accidents involving airline pilots was the "No Pre-Flight WX Briefing OR Intentional Flight Operations in Instrument Conditions." However, in statistical testing, the proportions corresponding to this violation category for the two groups of accident pilots were comparable ($p > 0.005$).

4. Discussion and conclusions

We show herein that, for general aviation operations, airline pilots show both safety improvements and deficits relative to PPL-IFR airmen. Regarding improvements the absence of in-flight loss-of-control accidents in degraded visibility was notable for airline pilots. Conversely, and initially surprising, these airmen were more likely to experience a ground loss of directional control during the landing roll. Finally, in regard to violations of the FAA regulations, despite the regimented nature of air carrier operations, airline pilots showed a greater proclivity for disregarding the minimum altitudes prescribed by 14CFR 91.

The safety of the airline pilots operating in degraded visibility, as witnessed by an absence of any in-flight loss-of-control accidents, merits some discussion especially since such mishaps under corresponding conditions in the general aviation sector are frequent and moreover carry a high fatality rate (National Transportation Safety Board, 2018). Certainly, these professional airmen have a high exposure to such weather conditions as part of their professional occupation. In contrast PPL-IFR pilots eschew operating in such weather (Federal Aviation Administration, 2015) and struggle to maintain currency to legally operate in instrument conditions (Weislogel, 1983). Nevertheless, transport-category aircraft are highly automated and there is current debate as to whether such automation adversely affects stick-and-rudder skills. Indeed, in a study of Boeing 747 aircraft pilots tasked with performing an instrument approach in which aircraft automation was progressively degraded (Casner et al., 2014), 44% were in error in identifying the missed approach fix and 16% descended below the minimum altitudes prescribed by the approach chart. How then do these findings reconcile with the stellar performance of airline pilots operating light aircraft with less automation (Federal Aviation Administration, 2015) in degraded visibility in the general aviation environment? We suspect that a combination of increased experience operating transport-category aircraft under such conditions and ATP certification (Boyd & Peters, 2015; Li, Baker, Grabowski, & Rebok, 2001) demanding a higher level of proficiency in instrument flight (relative to the IFR rating held by PPL airmen) more than offset any decrements caused by frequent automation usage.

The sub-classification of landing accidents related to pilot technique informs performance deficiencies for both the PPL-IFR and airline pilot cohorts. Indeed, the preponderance of landing accidents caused by poor landing speed control (mainly high-energy) for the former airmen contrasting with the absence of such mishaps for the latter pilots is noteworthy. Our findings are congruent with those of prior studies (Boyd, 2019; Goode et al., 1976) reporting on the tendency of general aviation pilots to carry excessive

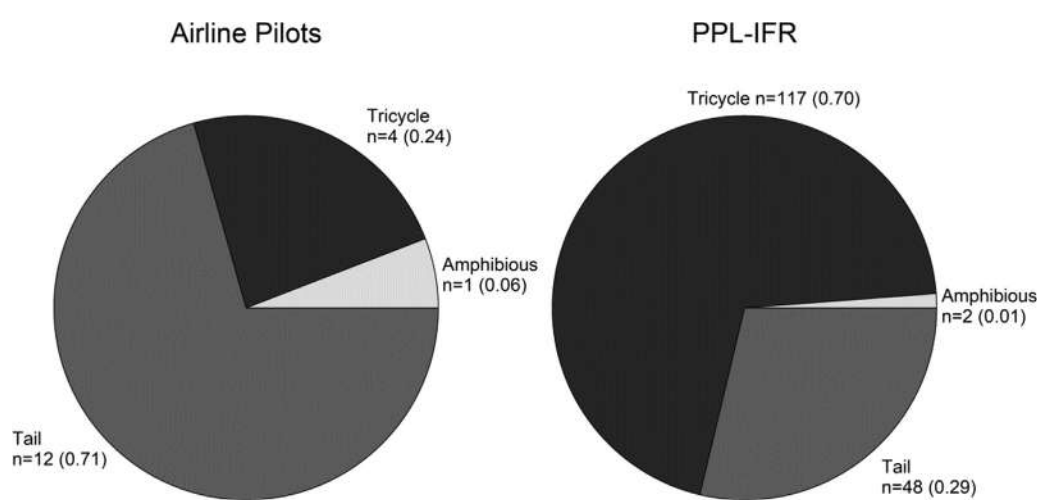


Fig. 4. Landing gear type for accidents ascribed to a deficiency in pilot technique. The proportion of accident aircraft with the specified undercarriage equipment shown (values in parentheses) for landing mishaps ascribed to a deficiency in pilot stick-and-rudder skills. Differences in proportions was tested using a Fisher's Exact Test. *n*, accident count.

Table 2
Categories of FAA violations for airline and PPL-IFR pilots.

FAA violation	Airline Pilots		PPL-IFR		*P* value
	Count (*n*)	Fraction	Count (*n*)	Fraction	
Disqualifying Medical Condition/Use of Illegal Drugs	0	0.00	4	0.10	>0.05
Intentional Visual Flight Departure into Instrument Conditions OR No Pre-Flight WX Briefing	3	0.43	22	0.54	>0.05
Lack of IFR Currency	0	0.00	2	0.05	>0.05
Maneuvering Flight below Legal Minimum Altitude	4	0.57	7	0.17	0.021
Un-Airworthy aircraft	0	0.00	6	0.15	>0.05
					>0.05
TOTAL	7	1	41	1	

The count (*n*) and proportion (fraction) of accidents in which the NTSB cited the specified FAA violation is tabulated. P values were derived from adjusted residuals from a Fisher's Exact Test. Wx, weather.

landing speeds (higher than V-$_{Ref}$ –airplane speed in the landing configuration, at the point where it descends through the 50 ft. height) (Saini, 2010). Such a practice with transport category aircraft would lead to an abundance of runway overruns and air carrier pilots must adhere closely to the approach speed regimen. On the other hand, airline pilots relative to their PPL-IFR counterparts showed a greater deficiency in maintaining ground directional control during the landing roll. We argue that several reasons likely underlie this observation. First, airline pilots accrued a lower amount of time-in-type as evident from both a prospective survey of the airline pilot population-at-large as well as that for the accident airmen (median make-model flight times-132 and 261 hours for airline and PPL-IFR respectively). Second, compared with operating a transport category aircraft, light aircraft demand more control inputs for identical landing wind conditions. Third, and likely most important, is the preference of airline pilots for operating light aircraft with tail-wheel landing gear (conventional undercarriage). It is well established that such airplanes show ground handling characteristics at variance with tricycle aircraft (Federal Aviation Administration, 2016). Specifically, conventional aircraft are inherently unstable on the ground and exhibit an exaggerated tendency to weathervane during ground operations in a crosswind (Federal Aviation Administration, 2016). In regard to this latter point, we considered the possibility that the involved conventional under-carriage aircraft had unique ground handling characteristics based on: (a) being of experimental build or (b) less rigorous certification standards in effect for older aircraft. However, these arguments are unlikely for two reasons. First, none of the ground loss of directional control mishaps involved experimental (i.e., non-certificated) aircraft. Second, while indeed the involved aircraft were of older vintage and subject to earlier certification regulations (i.e., civil air regulations-CARs (Civil Aeronautics Board, 1949), such standards with respect to ground handling were identical to those promulgated for later aircraft certification per 14CFR 23.231-233 (Electronic Code of Federal Regulation, 2017) effective up to 2017.

Surprisingly, airline pilots involved in accidents did not show greater compliance with the FAA regulations than PPL-IFR airmen. For this accident category, more than half of mishaps were due to these airmen operating the aircraft below the minimum altitudes prescribed by 14CFR 91.119 (500 and 1,000 feet above ground for other-than-congested and congested areas respectively) (Electronic Code of Federal Regulation, 2015). Why is this? One must consider that airline operations are all conducted under IFR rules requiring adherence to minimum altitudes defined by jet routes, standard arrivals and departures (Federal Aviation Administration, 2015) absent for VFR operations. Whether airline pilots were unfamiliar with the minimum altitudes for VFR operations per 14CFR 91 (Electronic Code of Federal Regulation, 2015) or were deliberately operating contrary to such regulations is currently unknown. Based on anecdotal information we suspect the former. Thus, for three of the four minimum altitude infractions, in their NTSB statements one pilot admitted to flying " along a creek" another, "through a valley" with the third airman stating descending to what he "thought was a safe VFR altitude." Notably, none of these accidents were due to degraded visibility ruling out "scud-running" as a causal factor. Another question raised by this infraction relates to the role of surveillance evident in the airline industry but absent from general aviation. Consequently, general aviation pilots may be tempted to infringe such minimum altitudes with immunity nevertheless developing a greater skillset with respect to operating below legal minimum altitudes.

Also noteworthy was the disregard by airline pilots for the FAA regulations necessitating preflight weather briefings and intentional flight into instrument conditions. We entertain the possibility that the former transgression relates to the role of the airline dispatcher in preparation of a weather briefing for their pilots. It may be that: (a) the airline pilot is so habituated to receiving this prepared material that such a task is overlooked for general aviation and/or (b) he/she may be unaware of the tools to obtain a weather brief via official sources typically used by the general aviation pilot.

Although our research is the first to report on airline pilot safety in general aviation, an older study of accidents spanning the 1973–1983 period merits discussion (Salvatore, Stearns, Huntley, & Mengert, 1986). The authors of that report noted that most ATP-certificated pilot accidents were due to aerobatics whereas, in the current study, aerobatics was cited for a single airline pilot accident. Moreover, only 4% of the airline pilot population-at-large survey respondents indicated this as the primary purpose of their general aviation flights. How can the differences in the results between the two studies be reconciled? A key difference in study design is pertinent. Specifically, the Salvatore and co-author study was not limited to airline pilots per their two inclusion criteria: (a) ATP-certificated and (b) a self-description as a "professional pilot." Thus, the cohort would also include pilots engaged in charter operations (14CFR 135), corporate flying and other non-air carrier professions with corresponding lower levels of training/recurrency/oversight. In addition, much has changed in general aviation over the intervening three decades in regard to technology such as in-flight data-linked weather and in the case of general aviation scenario-based training (Federal Aviation Administration, 2017).

Our study was not without limitations. First, the absence of denominator data for both pilot cohorts operating under 14CFR 91 regulations precluded the determination of accident rates. Second, the count of airline pilot accidents was, in some cases, small. Third, risk exposure was determined in a prospective study with accident data obtained in a retrospective query. Fourth, type rating data, used as one of the criteria to operationally define an airline pilot, was in some instances based on information current at the time (2019–2020) over which the research was conducted. As a result, for a subset of non-fatal accident pilots, a type rating may have been achieved after the mishap. Fifth, we accept that the multiple criteria used concurrently (ATP certification, a Class 1 Medical and type rating in a transport category aircraft) to operationally define an airline pilot might also lead to the inclusion of a few airmen who do not fly for an air carrier. Finally, (and not addressed in the current study) it would be of particular interest in a future survey to determine how airline and PPL-IFR pilots' views compare with respect to safe operations of a light aircraft. In a similar vein, endeavors to capture accident pilot attitudes in NTSB reports with respect to "thrill-seeking" in an environment absent for surveillance are lacking.

Although the objective of the current study was to determine the safety of airline pilots in the general aviation environment, the findings inform performance deficiencies for both these and PPL-IFR pilots which warrant redress. Notably, regarding the preponderance of in-flight loss-of-control fatal accidents involving PPL-IFR pilots, such airman would be well served by increasing the frequency of recovery from unusual attitudes maneuvers by reference to instruments in recurrency training. Moreover, for airmen with deficient IFR proficiency skills, safety could be improved by development of computer-based training systems which provide pilots with skills to recognize cues (e.g., cloud bases, visibility, darkening) associated with impending IMC as reported elsewhere (Wiggins & O'Hare, 2003). The wide availability of advanced aviation training devices should make for a cost-effective means of achieving/maintaining such proficiency. PPL-IFR safety would also benefit from an increased emphasis on landing energy/speed management in training/recurrency. As to airline pilot safety, airmen seeking to operate a light aircraft with tail-wheel landing gear should consider, post tail-wheel endorsement, additional dual time with an instructor (well experienced in conventional landing gear operations) focusing on landing/ground operations particularly under crosswind conditions. This recommendation would be on par with the initial operating experience required (14CFR 121.913) for airline pilots (Electronic Code of Federal Regulation, 2017). Finally, it would behoove airline pilots to adhere more closely (and if necessary familiarize themselves with) to 14CFR 91 regulations pertinent to general aviation operations (in particular minimum altitudes) towards improving their safety whilst operating light aircraft.

5. Practical applications

For PPL-IFR airmen, training/recurrency should focus on unusual attitude recovery and managing approach speeds. Airline pilots should seek additional instructional time regarding landing tail-wheel aircraft and become familiar with 14CFR 91 rules covering minimum altitudes. Lastly, future accident reporting should seek to capture airline pilot attitudes in the "overconfidence/misplaced motivation" nano-codes in the Preconditions for Unsafe Acts/Adverse Mental States domain per the established Human Factor Classification System (Shappell & Wiegmann, 2001; Shappell, Detwiler, Holcomb, Hackworth, Boquet, & Wiegmann, 2006).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We express our gratitude to Drs. Anthony Tvaryanas, Christy Hileman and Joe Mooney (FAA Civil Aerospace Medical Institute) for providing type ratings for some of the accident pilots. We are also thankful to AvWeb Flash, Embry Riddle Alumni Association, JetWhine, AOPA E-Digest and the Southwest Pilot Forum for advertising the web link to the online pilot survey.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

References

Agresti, A. (2012). *Categorical data analysis* (3rd ed.). Hoboken: Wiley.
Aviation DB. (2019). AviationDB Aircraft Query. http://www. aviationdb. com/Aviation/AircraftQuery. shtm.
Boyd, D. D. (2017). A review of general aviation safety (1984-2017). *Aerospace Medicine and Human Performance, 88*, 657–664.
Boyd, D. D. (2019). Occupant injury severity in general aviation accidents involving excessive landing airspeed. *Aerospace Medicine and Human Performance, 90*, 355–361.
Boyd, D. D., & Peters, C. (2015). Should charity air medical organizations require commercial certification of their pilots. *Air Medical Journal, 34*, 188–194.
Casner, S. M., Geven, R. W., Recker, M. P., & Schooler, J. W. (2014). The retention of manual flying skills in the automated cockpit. *Human Factors, 56*, 1506–1516.
Cessna. (2020). Cessna Skyhawk specifications. http://cessna. txtav. com/en/single-engine/skyhawk, http://cessna.txtav.com/en/single-engine/skyhawk
Civil Aeronautics Board. (1949). Airplane airworthiness-normal, utility, acrobatic and restricted purpose categories. http://www. navioneer. org/riprelay/Yet%20More%20Navion%20Files/car_part3. pdf.
Electronic Code of Federal Regulation. (2015). Operating requirements: Domestic, flag, and supplemental operations: INITIAL, transition and recurrent training and checking requirements. https://www.ecfr.gov/cgi-bin/text-idx?SID=913cace5e186609a4b2e48a8474f467b&mc=true&node=pt14.3.121&rgn=div5#se14. 3. 121_1414.
Electronic Code of Federal Regulation. (2012). Airworthiness standards: Normal, utility, acrobatic and commuter category airplanes. http://www. ecfr. gov/cgi-bin/text-idx?SID=5ffea7e4489b0113fefc117f1b9fc96a&mc=true&node=pt14. 1. 23&rgn=div5#se14. 1. 23_11.
Electronic Code of Federal Regulation. (2015). General operating and flight rules. http://www.ecfr.gov/cgi-bin/text-idx?node=14:2.0.1.3.10.
Electronic Code of Federal Regulation. (2017). Airworthiness standards: Transport category airplanes. https://www. ecfr. gov/cgi-bin/text-idx?SID=ebee64813226fcbca312d2fc7319d0ff&mc=true&node=pt14. 1. 25&rgn=div5.
Electronic Code of Federal Regulation. (2017). Airworthiness standards: Normal, utility, acrobatic and commuter category airplanes. http://www. ecfr. gov/cgi-bin/text-idx?SID=5ffea7e4489b0113fefc117f1b9fc96a&mc=true&node=pt14. 1. 23&rgn=div5#se14. 1. 23_11.
Electronic Code of Federal Regulation. (2017). Airworthiness standards: Normal, utility, acrobatic and commuter category airplanes. https://www. govinfo. gov/content/pkg/CFR-2017-title14-vol1/pdf/CFR-2017-title14-vol1-chapI-subchapC. Pdf.
Electronic Code of Federal Regulation. (2017). Operating requirements: Domestic, flag and supplemental operations. http://www.ecfr.gov/cgi-bin/text-idx?node=14:2.0.1.3.10, http://www.ecfr.gov/.
Electronic Code of Federal Regulation. (2018). Certification: Pilots, flight instructors, and ground instructors. https://www.ecfr.gov/cgi-bin/text-idx?SID=ff99c129f19bfc12ab36a66da85735d5&mc=true&node=se14.2.61_156&rgn=div8.
Electronic Code of Federal Regulation. (2020). Aeronautical experience: Airplane category rating. https://www.ecfr.gov/cgi-bin/retrieveECFR?gp=&r=PART&n=14y2. 0. 1. 1. 2#se14. 2. 61_15.
Federal Aviation Administration (2008). *On landings Part II*. Oklahoma City: U.S. Department of Transportation.
Federal Aviation Administration. (2004). Flight operational quality assurance. AC 120-824. 1-8.
Federal Aviation Administration. (2012). Threat and error management (TEM) and line operations safety assessment (LOSA) Training. https://www. faa. gov/about/initiatives/maintenance_hf/losa/training/media/full_awareness_2_2012-cami_v3. pdf.
Federal Aviation Administration. (2014). Standard operating procedures and pilot monitoring duties for flight deck crewmembers. AC 120-71B.
Federal Aviation Administration. (2015) U.S. Civil Airmen Statistics. https://www.faa.gov/data_research/aviation_data_statistics/civil_airmen_statistics/.
Federal Aviation Administration. (2015). Terminal procedures. http://www. faa. gov/air_traffic/flight_info/aeronav/digital_products/dtpp.
Federal Aviation Administration. (2015). General Aviation and Part 135 Activity Surveys. http://www. faa. gov/data_research/aviation_data_statistics/general_aviation.
Federal Aviation Administration. (2016). Aircraft Registry (FAA). http://www.faa.gov/licenses_certificates/aircraft_certification/aircraft_registry/releasable_aircraft_download.
Federal Aviation Administration. (2016). Safety Management System. 8000.369B. 1–14.

Federal Aviation Administration. Transition to Tailwheel Airplanes. In: Airplane Flying Handbook (FAA-H-8083-3B), Oklahoma City:U.S. Department of Transportation, 2016, p. 1-8.
Federal Aviation Administration. (2017). Introduction of scenario-based training. https://www. faasafety. gov/files/gslac/library/documents/2007/Sep/19529/Introduction%20to%20Scenario-Based%20Training. pdf.
Federal Aviation Administration. (2017) Airport Data and Contact Information. https://www. faa. gov/airports/airport_safety/airportdata_5010/.
Federal Aviation Administration. (2018). Instrument rating-airplane. Airman Certification Standards. FAA-S-ACS-8B.
Federal Aviation Administration. (2018). Aeronautical Information Manual, Federal Aviation Administration, 2017 Washington, DC, 1-16-1-17.
Federal Aviation Administration. Airline Transport Pilot and Type Rating for Airplane. Airman Certification Standards. 2019, FAA-S-ACS-11
Federal Aviation Administration. (2019). Airmen Certification Database. https://www. faa. gov/licenses_certificates/airmen_certification/releasable_airmen_download/.
Field, A. (2009). *Discovering statistics using IBM SPSS statistics*. Thousand Oaks, California: SAGE Publications.
Goode M. W., O'Bryan, T. C., Yenni, K. R., Cannaday, R. L., & Mayo, M. H. (1976). Landing practices of general aviation pilots in single-engine light airplanes. TN D-8283.
International Air Transport Association. (2018). Loss of Control In-flight (LOC-I). https://www. iata. org/whatwedo/safety/Pages/loss-of-control-inflight. Aspx.
Kirby, J. (2020). Social acceptance of increased usage of the ballistic parachute system in a general aviation aircraft. *Aerospace Medicine and Human Performance, 91*, 86–90.
Li, G., & Baker, S. P. (2007). Crash risk in general aviation. *JAMA, 297*, 1596–1598.
Li, G., Baker, S. P., Grabowski, J. G., & Rebok, G. W. (2001). Factors associated with pilot error in aviation crashes. *Aviation Space and Environment Medicine, 72*, 52–58.
Modern Airliners. (2020). Boeing 737 Specifications. https://modernairliners. com/boeing-737/boeing-737-specifications/#Specs%20737%20Next%20gen%20and%20Max.
National Transportation Safety Board. (2015). NTSB Accident Database. http://app. ntsb. gov/avdata/Access/.
National Transportation Safety Board. (2016). Most wanted list of transportation safety improvements. https://www. ntsb. gov/safety/mwl/Documents/MWL2016_Brochure_web. pdf.
National Transportation Safety Board. (2018). General aviation accident statistics. https://www. ntsb. gov/news/speeches/EWeener/Documents/weener-20181215. pdf.
Saini, M. (2010). Flying by the numbers. A precise approach to landing. *FAA Safety Briefing*, 11–14.
Salvatore, S., Stearns, M. D., Huntley, M. S., & Mengert, P. (1986). Air transport pilot involvement in general aviation accidents. *Ergonomics, 29*, 1455–1467.
Shappell S. A., Detwiler, C., Holcomb, K. A., Hackworth, C., Boquet, A., & Wiegmann, D. A. (2006). Human error and commercial aviation accidents: A comprehensive, fine-grained analysis using HFACS. DOT/FAA/AM-06/18.
Shappell, S. A., & Wiegmann, D. A. (2001). Applying reason: The human factors and classification system (HFACS). *Human Factors and Aerospace Safety, 1*, 59–86.
Weislogel, G. S. (1983). *Study to determine the IFR operational profile and problems of the general aviation single pilot*. National Aeronautics and Space Administration. *NASA-CR-3576, NAS 1.26:3576*.
Wiggins, M., & O'Hare, D. (2003). Weatherwise-evaluation of a cue-based training approach for the recognition of deteriorating weather conditions during flight. *Human Factors, 45*, 337–345.

Douglas Boyd Ph.D, a research scientist and commercial-certificated Citation type-rated pilot has published 90 peer-reviewed papers (18 aviation-related). He received the Aerospace Medical Association (AsMA) Joseph L Haley and Arnold Tuttle awards for two papers on accident injury severity in helicopter emergency medical services and in-flight general aviation pilot decision-making with regard to extreme convection. Dr. Boyd serves as on the AsMA safety and the General Aviation (GA) Joint Steering committees. He serves as a FAA Safety Team (FAAST) representative, has made 8 presentations in this program and penned 15 aviation magazine articles to disseminate research findings to GA pilots.

Mark T. Scharf, PhD, an Assistant Professor with Embry-Riddle Aeronautical University-Worldwide, has an aviation background including military aviation (USA F-16, F15, and T-37 trainer), major airline duties with type ratings in the B-727, B-737, B-757/767, and the B-777, as well as extensive general aviation experience including ownership of a Piper Twin Comanche and Piper Archer II. In addition to his Air Transport Pilot certificate, he is also a Certificated Flight Instructor, Flight Instructor (Instruments) and Multi-Engine Instructor. His academic background includes publications in the areas of technology acceptance, crew resource management, safety management systems, and pilot attitudes toward safety. His current research emphasis (2 presentations and 3 publications) is in the area of developing a taxonomy for commercial pilot operations safety.

David Cross, PhD, an Assistant Professor with Embry-Riddle Aeronautical University-Worldwide, has an extensive aviation background including piloting the A-320, B-737, B-757/767, B-777, and B-787 for a major U. S airline. His academic background includes numerous publications in the areas of technology acceptance, crew resource management, safety management systems, and pilot attitudes toward safety. His current research emphasis (2 presentations and 3 publications) is in the area of developing a taxonomy for success in commercial pilot operations.

Journal of Safety Research 76 (2021) 238–247

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

A data mining approach to deriving safety policy implications for taxi drivers

Jiwon Park [a], Seolyoung Lee [b], Cheol Oh [c,*], Byongho Choe [d]

[a] Department of Smart City Engineering, Hanyang University at Ansan, 55 Hanyangdaehak-ro, Sangnok-gu, Ansan-city, Gyeonggi-do 15588, Republic of Korea
[b] Department of Smart City Research, Seoul Institute of Technology Principal Researcher, 8F, Maebongsan-ro 37, Mapo-gu, Seoul, 03909, Republic of Korea
[c] Department of Transportation and Logistics Engineering, Hanyang University at Ansan, 55 Hanyangdaehak-ro, Sangnok-gu, Ansan-city, Gyeonggi-do 15588, Republic of Korea
[d] Transportation Safety Research and Development Institute, Korea Transportation Safety Authority, 77 Hyeoksin 8-ro, Gimcheon-city, Gyeongsangbuk-do 39660, Republic of Korea

ARTICLE INFO

Article history:
Received 11 June 2020
Received in revised form 11 September 2020
Accepted 22 December 2020
Available online 7 January 2021

Keywords:
Traffic safety
Taxi crashes
Random forest method
Policy implications
Data mining

ABSTRACT

Introduction: Traffic safety issues associated with taxis are important because the frequency of taxi crashes is significantly higher than that of other vehicle types. The purpose of this study is to derive safety implications to be used for developing policies to enhance taxi safety based on analyzing intrinsic characteristics underlying the cause of traffic accidents. *Method:* An in-depth questionnaire survey was conducted to collect a set of useful data representing the intrinsic characteristics. A total of 781 corporate taxi drivers participated in the survey in Korea. The proposed analysis methodology consists of two-stage data mining techniques, including a random forest method, with data that represents the working condition and welfare environment of taxi drivers. In the first stage, the drivers' intrinsic characteristics were derived to classify four types of taxi drivers: unspecified normal, work-life balanced, overstressed, and work-oriented. Next, priority was determined for classifying high-risk taxi drivers based on factors derived from the first analysis. *Results:* The derived policies can be categorized into three groups: 'the development of new policies,' 'the improvement of existing policies,' and 'the elimination of negative factors.' Establishing a driving capability evaluation system for elderly drivers, developing mental health management programs for taxi drivers, and inspecting the taxi's internal conditions were proposed as new policies. Improving the driver's wage system, supporting the improvement of rest facilities, and supporting the installation of security devices for protecting taxi drivers are methods for improving existing policies to reinforce the traffic safety of taxi drivers. Last, restricting overtime work for taxi drivers was proposed as a policy to eliminate negative factors for improving taxi traffic safety. *Practical Applications:* It is expected that by devising effective policies using the policy implications suggested in this study, taxi traffic accidents can be prevented and the quality of life of taxi drivers can be improved.

1. Introduction

Compared to traffic accidents in commercial vehicles and noncommercial vehicles, traffic accidents associated with commercial vehicles are relatively severe. According to the traffic accident statistics by the Korean Traffic Accident Analysis System (TAAS), the reduction rate of the traffic accident fatality of noncommercial vehicles from 2009 to 2018 was approximately 39%, while the reduction rate of traffic accident fatality for commercial vehicles in the last 10 years was approximately 31%, which is relatively lower than the former case. In addition, according to the traffic accident statistics of the TAAS, from 2014 to 2018, taxis occupied approximately 46% of traffic accidents, which was relatively higher than other modes such as buses (17.5%), rental cars (15.4%), and trucks (12.8%). In addition, the number of traffic accidents in taxis increased by 1.46% in 2018 compared to 2017, and the number of injured persons also increased by 0.03%. Therefore, it is necessary to prepare for effective countermeasures based on an in-depth analysis of the causes of taxi-related accidents to enhance traffic safety on the road.

The main causes of traffic accidents can be categorized into three classes: road environmental factors, vehicle factors, and driver factors. As a result of analyzing how much road, vehicle, and driver factors contribute to traffic accidents, road factors account for 28–34%, vehicle factors account for 8–12%, and driver factors account for 93–94% (Do, 2013). In this context, a thorough

* Corresponding author.
E-mail addresses: wldnjsdm@hanyang.ac.kr (J. Park), lsy0717@hanyang.ac.kr (S. Lee), cheolo@hanyang.ac.kr (C. Oh), byongho.choe@kotsa.or.kr (B. Choe).

https://doi.org/10.1016/j.jsr.2020.12.017
0022-4375/

understanding of drivers' intrinsic and extrinsic characteristics that potentially lead to high accident risks is fundamental to preventing accidents and reducing the severity, which motivates our study.

This study focuses on identifying the intrinsic characteristics of taxi drivers in terms of traffic safety. The intrinsic characteristics include demographics, living and working environments, and health conditions in this study. An in-depth questionnaire survey was conducted to collect a set of useful data representing the intrinsic characteristics. A total of 781 corporate taxi drivers participated in the survey. Then, data obtained from the survey were incorporated into corresponding individual taxi drivers' accident-related data archived from the commercial taxi driver management system operated by the Korean Transportation Safety Authority (KOTSA). A two-stage data mining approach based on a random forest (RF) method was conducted to identify factors affecting taxi accidents. A notable feature of the proposed methodology was to extract factors that represent the characteristics by the type of taxi drivers and then to derive the priority of extracted factors capable of explaining traffic safety effectively. The type of taxi drivers was classified into four groups, unspecified normal, work-life balanced, overstressed, and work-oriented, which were defined by the literature (Commercial Vehicle In-depth Accident Study: Corporate Taxi, 2019). In the first stage, factors capable of distinguishing the type of taxi driver were extracted. Next, meaningful factors capable of identifying high-risk taxi drivers were prioritized by the second RF analysis. Last, this study derived a set of policy implications for the prevention of taxi crashes from three perspectives: personal, working, and health and welfare. The derived policies can be divided into the development of new policies, the improvement of existing policies, and the elimination of negative factors. In addition, the main purpose and key issues of each policy are discussed.

2. Literature reviews

The literature review in this study was conducted to study the relationship between the characteristics of taxi drivers and traffic accidents, and to study random forest methodology. As a result of the literature reviews, it was found that understanding the internal characteristics of taxi drivers is essential to analyzing traffic accidents in taxis. In addition, RF methods can be effectively used for identifying significant factors contributing to traffic accidents.

2.1. Characteristics of taxi drivers in traffic safety

Vahedi et al. (2018) showed that factors such as the age of taxi drivers, marital status, violation of laws and regulations, and the amount of daily traffic can affect traffic accidents in taxis. As a result of research on the causes and factors of traffic accidents of taxi drivers, Al-Ghamdi (2000) found that taxi drivers' safety distance violations and speeding are the main causes of taxi traffic accidents. Borowsky and Oron-Gilad (2013) conducted a study on the effect of driving experience on risk perception ability for drivers with low driving experience, drivers with high driving experience, young drivers, and taxi drivers. The study found that taxi drivers were more sensitive to danger than other drivers. Maag et al. (1997) conducted a study on the relationship between the taxi driver's visual ability and the frequency of accidents. As a result, the taxi driver's accident frequency was higher than that of the general driver, and specifically, a driver's visual deterioration was more frequent. Li et al. (2019) suggested that taxi drivers with a high risk of accidents tend to have a longer average daily driving time and less daily rest time. Meng et al. (2019) analyzed the effect of fatigue on the driving ability of taxi drivers. As a result of the analysis, it was found that taxi driver fatigue had an effect on brake control, lane control, and steering control. Apantaku et al. (2012) conducted a study on the work and health characteristics of taxi drivers in Chicago, Illinois, USA, and found that cancer screening participation, exercise, and fruit and vegetable consumption rates were low compared to general drivers. Lim and Chia (2015) conducted a study on the health characteristics of taxi drivers. As a result of the study, it was found that obesity, hypertension, and diabetes were the highest among taxi driver diseases, and taxi drivers experienced high fatigue due to lack of sleep. Miyamoto et al. (2008) suggested the limited working space in a taxi and whole-body vibration that occurs during operation as a cause of low back pain for taxi drivers. Bulduk et al. (2014) suggested that taxi drivers have very high exposure to musculoskeletal disorders, which is significantly related to lack of rest, age, and career. Chen et al. (2005) found that a taxi driver may have a higher risk of developing cardiovascular disease in a month with more travel time. Hansen et al. (1998) found that taxi drivers may increase the risk of cancer by inhaling carcinogens while on duty.

2.2. Random forest method for traffic safety analysis

Ma and Cheng (2016) presented that the random forest method can be used to analyze a large number of explanatory variables without losing data. Lee et al. (2019) analyzed spatiotemporal data by processing weather and road condition information to predict road conditions using a RF method with 95% accuracy. Ximiao Jiang et al. (2016) used a large quantity of data to derive the influencing factors of traffic accidents according to the type of accident for each section using an RF method. Siddiqui et al. (2012) used decision tree and random forest techniques to derive essential variables related to the severity of collisions for each traffic analysis zone in four areas of Florida. Abdel-Aty et al. (2008) used an RF technique to evaluate traffic safety on Dutch highways and suggested that the random forest technique was a relatively effective classifier compared to decision trees. Harb et al. (2009) used a random forest technique to derive important factors related to drivers traffic accident avoidance behavior.

2.3. Research opportunity

Many researchers have studied the relationship between the characteristics of taxi drivers and traffic accidents. However, there is little research to address work, welfare, life, and health-related characteristics along with traffic accident data. This study attempts to investigate the causes of traffic accidents for corporate taxis in terms of the internal characteristics of drivers. Historical accident data, commercial vehicle driver aptitude test data, and in-depth cause survey data are used to identify factors affecting taxi driver safety. In addition, this study differs from previous research in that it suggests the necessity of legal and institutional management to reduce traffic accidents in taxi and reduce accident risk group classification by proposing traffic safety policy implications for taxi drivers.

3. Data

3.1. Data collection

The dataset used in this study was established by two different resources, including an in-depth questionnaire survey for identifying intrinsic driver factors potentially leading to crashes and commercial vehicle driver management system (CVDMS) data operated by KOTSA. More details for the data are presented in the subsequent sections.

3.2. In-depth investigation of corporate taxi drivers

A survey to identify potential factors leading to crash occurrence was conducted as a part of this study. The investigation of the causes of traffic accidents in corporate taxis was conducted from September to October 2019. KOTSA conducts traffic safety inspections annually for companies with a history of traffic crashes in which one fatal or three seriously injured persons occurred. A total of 89 taxi companies participated in the survey, among the companies that were inspected between the 3rd quarter of 2018 and the 2nd quarter of 2019, participated in the survey of this study. In-depth interviews using questionnaires were conducted for taxi drivers and a safety officer of each taxi company. A face-to-face survey, which took approximately 20 minutes for each subject, was conducted with 781 taxi drivers and 89 safety managers. Fig. 1 presents the distribution of taxi drivers by administrative district based on the address of the taxi company participating in the survey. The major regions of Korea were divided into the Seoul metropolitan area, Yeongnam area, Chungcheong area, Honam, and Gangwon area. The Gangwon area was excluded from the survey as there were neither fatal nor seriously-injured crashes. The occupancy rate of taxi drivers by each region is 48%, 31%, 19%, and 2%, respectively.

The questionnaire consisted of questions that included drivers' demographics, disease, mental and physical health conditions, working environment, economic situation, living conditions represented by interpersonal and family relationships, and leisure activities. A total of 230 candidate variables to be used for the proposed analysis method were extracted by processing the survey result by taking the characteristics of the questionnaire, such as multiple-response questions, into consideration. The questionnaire for the safety manager was focused on the management and training status of the drivers and the working environment of the taxi company. Fourteen variable candidates were extracted using questions related to driver welfare of taxi companies, such as whether to conduct safety training for elderly drivers and whether to have rest facilities.

3.2.1. Commercial vehicle driver management system

Two candidate variables were derived from the results of the driver aptitude test, which consists of two different types of tests, referred to as test I and test II in this study, conducted by the KOTSA. The aptitude test I is a driving adequacy test that is performed for all commercial vehicle drivers. Aptitude test II is conducted for drivers involved in a serious accident and drivers expected to be incapable of safe driving due to illness and other reasons. Both tests evaluate drivers in terms of driving ability based on 13 test items. In this study, the average grade for each driver obtained from test I and test II was calculated and used as a variable. The driving ability was considered to be excellent, as the calculated average grade was closer to 1.

3.3. Establishment of the dataset

3.3.1. Input variables

A set of candidate variables was obtained from the in-depth survey and CVDMS for inputs of the proposed analysis method. A data processing procedure as shown in Fig. 2 was devised to select meaningful variables to be used for conducting the random forest analysis. A total of 246 candidate variables evaluated by the proposed variable selection procedure consisting of two steps in terms of whether given hypotheses are accepted or not. The hypothesis defined in this study is that the average difference of a variable in cases where an accident does and does not occur is statistically significant, and a theoretical causal relationship is also feasible.

In Step 1, a total of 246 candidate variables were evaluated in terms of whether a given hypothesis is accepted or not. The hypothesis defined in this study is that the average difference of a variable in cases where an accident does and does not occur is statistically significant, and a theoretical causal relationship is also feasible. Regarding the continuous variables, the average difference was investigated by conducting a t-test according to whether the taxi driver experienced accidents in the last three years. Additionally, the chi-square test was adopted to identify the statistical significance of differences in nominal variables. The statistical

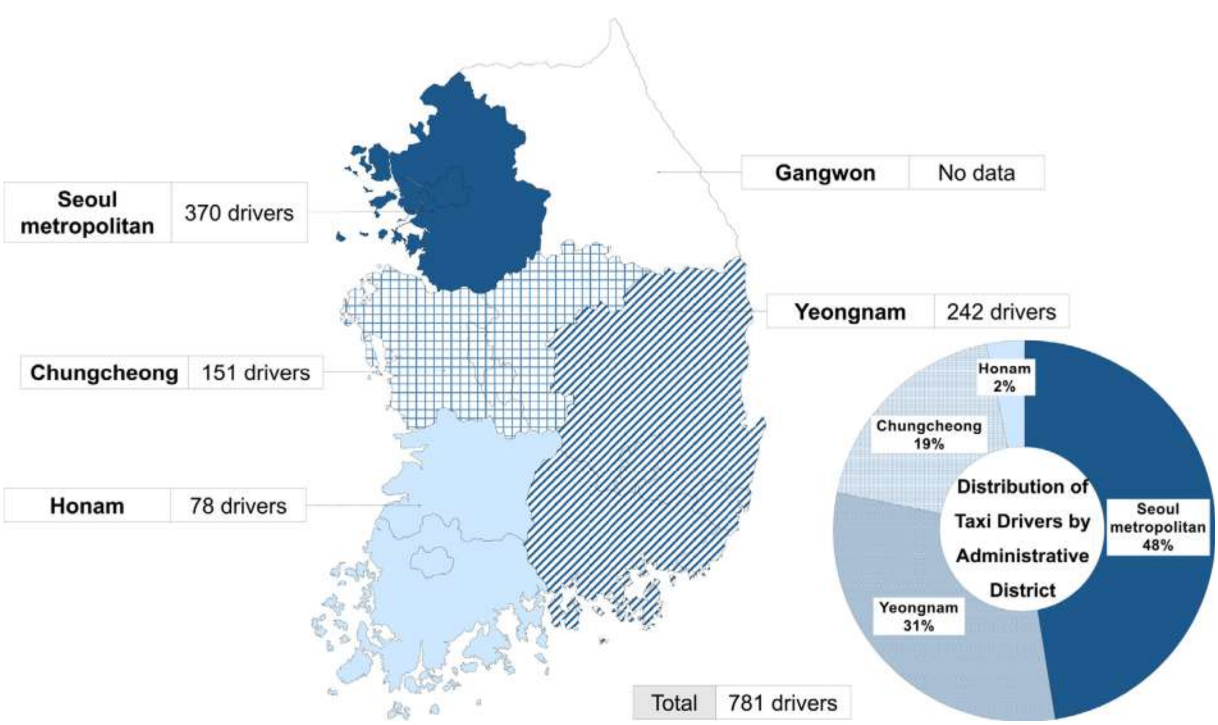


Fig. 1. Percentage of respondents by administrative region.

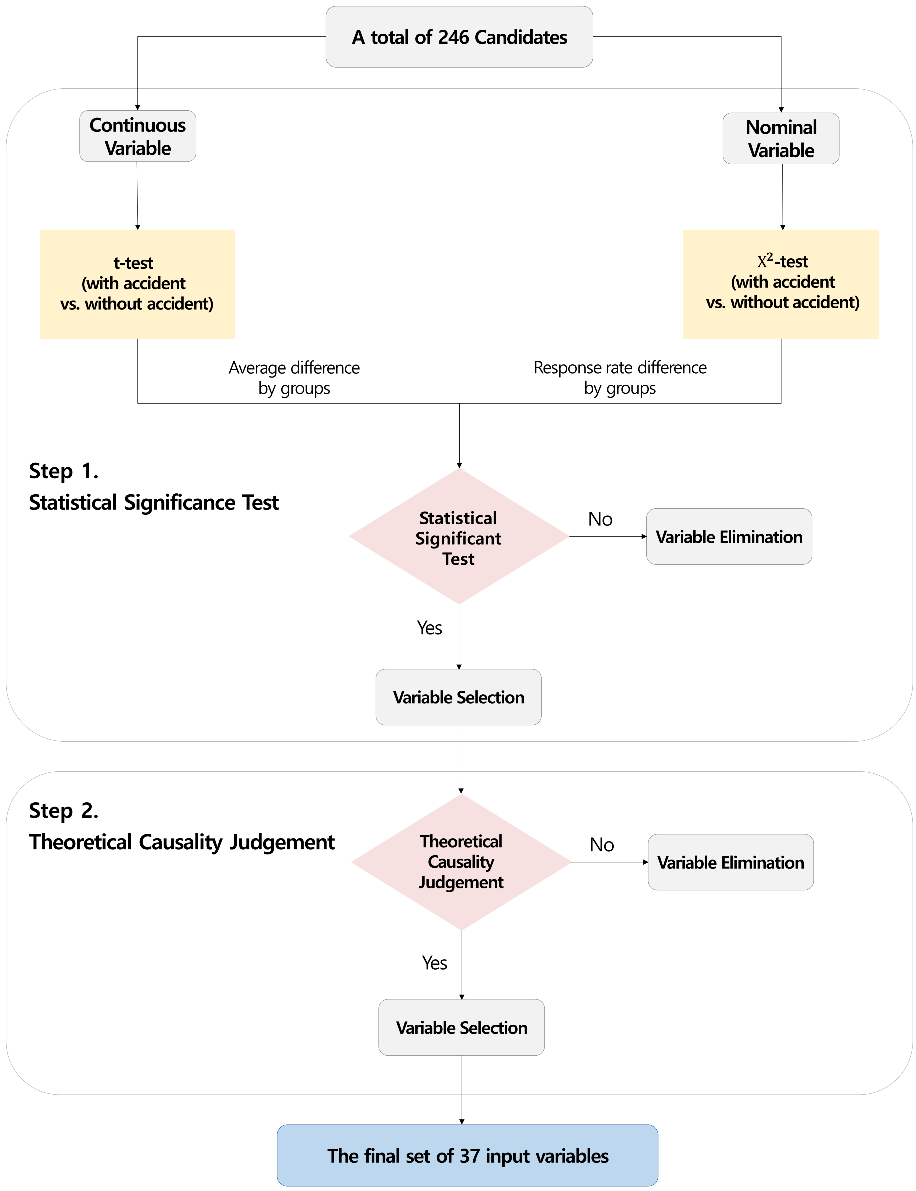


Fig. 2. Data processing procedure for input variables.

significance of the t-test and chi-square-test results was examined at a 90% confidence level. In Step 2, the test results were further evaluated to be the same as the hypothesis in light of the theoretical causation between the variables and traffic accidents. Variables were classified into three groups based on the test results, such as 'suitable,' 'unsuitable,' and 'unable to judge.' The suitable variable in this study indicates that there is a significant difference according to the occurrence of an accident, and the theoretical interpretation of the relationship is acceptable. Variables for which the hypothesis suitability cannot be judged refer to ones corresponding to a case where the ratio of nonresponse or biased-response is high. For example, drivers who conduct volunteer social services as a leisure activity cannot be obviously classified because the response rate of drivers with an accident experience is approximately 0.3%, and the response rate of drivers without an accident experience is 0%. It was identified by conducting the analysis in Step 1 that a total of 71 variables are suitable to be used for the inputs of the proposed methodology. Others that did not fit the hypothetical conditions were excluded from the set of input variables. Then, 37 taxi driver characteristic variables were selected as the final set of input variables for the random forest analysis by excluding variables with low theoretical causality from the 71 variables in Step 2.

3.3.2. Target variables

This study defines target variables to be used for the proposed two-stage random forest-based data mining method based on the taxi driver's types and accident history data. For the first random forest analysis, four target variables obtained from a previous study (Commercial In-Depth Accident Study (CIDAS) - Corporate Taxi, 2020) were adopted. Taxi drivers were classified into four groups: unspecified normal, work-life balanced, high-stress imbalanced, and work-oriented. Lee et al. (2020) classified taxi drivers into four groups through on a hybrid clustering methodology that performs hierarchical and non-hierarchical cluster analyses sequentially by utilizing internal characteristics of corporate taxi drivers. In categorizing taxi drivers, a total of 21 driver characteristics, which can be fallen into four classes including health, working, living, and crash conditions, were used for their analysis. Factors contributing to the classification of taxi drivers include waiting time, health care level, weekly working hours, crash severity, and so forth. The unspecified normal type is defined as the most common driver. The work-life balanced type is defined as one that is generally found to have good conditions in work, life, and heath. The high-stress imbalanced type is defined as one that is generally found to have bad conditions in work, life, and health. Finally, the work-oriented type is defined as one that is found to be relatively good in life satisfaction and health status, but one with high age and long working hours.

In the second random forest analysis, an accident risk group was defined and used as target variables for the classification based on the accident experience frequency and casualties per accident in the last three years. Taxi drivers who fall into the high-risk group have experienced at least one serious or fatal accident, and taxi drivers falling into the medium-risk group have experienced minor or injured traffic accidents in the last three years. Low-risk drivers are defined as those who have never experienced an accident.

In the first stage, factors capable of distinguishing the type of taxi driver were extracted. Next, meaningful factors capable of identifying high-risk taxi drivers were prioritized by the second RF analysis. Finally, policy implications to enhance taxi drivers' safety were derived based on the results of the proposed analysis method. Directions of policy development and the needs of institutional management of taxi drivers were discussed.

4. Methodology

The proposed methodology consisting of two-stage random forest analysis was devised to address the correlation between taxi driver intrinsic characteristics and traffic accidents as shown in Fig. 3. In the first stage, factors capable of distinguishing the type of taxi driver were extracted. It is important to extract variables that are capable of characterizing taxi driver types in that different types of drivers have different contributing factors potentially leading to accident occurrence. Next, meaningful factors, which were extracted from the first stage, capable of identifying high-risk taxi drivers were prioritized by the second RF analysis. Finally, policy implications to enhance taxi drivers' safety were derived based on the results of the proposed analysis method. Directions of policy development and the needs of institutional management of taxi drivers were discussed.

4.1. Random forest

Random forest is one of the ensemble machine learning techniques, and it is an advanced technique of decision tree analysis developed to address the problem of decision tree analysis (Breiman, 2001). In the random forest learning process, each tree is generated based on bootstrap samples that are randomly selected with replacement. The number of trees is set in advance

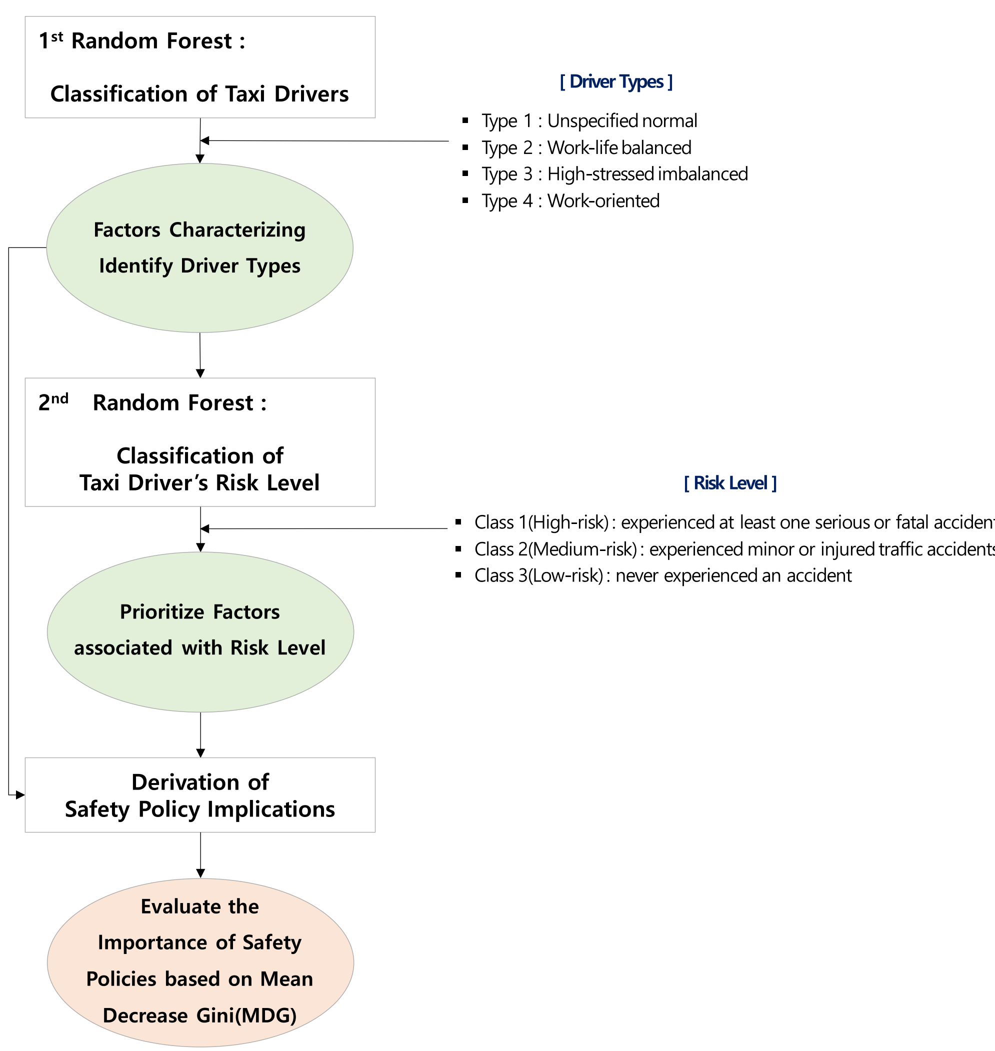


Fig. 3. Proposed Methodology.

by the analyst, and based on the results generated in each tree, the average value of the results for each tree is derived as the final prediction result. The random forest technique is known as a technique that can build a model with excellent predictive performance. This is because there is no restriction on variable selection so it is possible to prevent overfitting. The random forest learning process using bootstrap sampling proceeds in the following steps.

1) Generate Ts training datasets from the given training dataset through the bootstrap sampling.
2) Train a basic T sorter (tree).
3) Combine the basic sorter (tree) into one sorter (random forest).
4) Determine the final forecast by average and majority voting.

Observed values that are not used in the learning process of individual trees in the random forest are called out-of-bagging (OOB) and are used to validate the model. OOB is used to estimate the predicted probabilities by category and identify the variables that cause anomalies. The number of times that OOB is selected in all the decision trees of the random forest differs from each tree, and the predicted values are different for each tree. The probability of appropriately predicting the OOB observation of each observation belonging to category k in the original category is defined as follows.

$$\widehat{p_k}(x_i) = \frac{\sum j\in\overline{OOB_i} - I\left[\widehat{y}(x_i,t_j)=k\right]}{|OOB_i|}, \text{ for } k$$

I is an indicator function that indicates 1 when the value in the parentheses is true, 0 when the value is false, and $\widehat{y}(x_i, t_j)$represents the predicted category. $t_j$means the jth decision tree in the generated decision tree. OOB_i is a set of decision trees that is not used as an observed value in the learning process using bagging. Where a set of decision trees does not include x_i, the ratio of the number of decision trees predicting x_i to the k category is $\widehat{p_k}(x_i)$ (Deng et al., 2012). The verification method using OOB is as accurate as verification through new verification data (Breiman, 2001).

This study uses the mean decrease Gini (MDG) as an index to evaluate the significance of explanatory variables in random forests. MDG is the average of the GI index reductions for a particular explanatory variable in all trees. If the number of classification categories (i) is j, that is, $i = 1, 2, 3, \cdots, J$, the GI index is calculated as follows.

$$GI = \sum_{i=1}^{J} f_i(1 - f_i) = 1 - \sum_{i=1}^{J} f_i^2$$

$f_i$is the ratio of correctly classifying i to i category, and $1 - f_i$is the ratio of predicting i to another category. If the model perfectly classifies every category, f_i is 1, and the GI index is 0. The higher the MDG value of a particular variable indicates that this value is suitable for classifying a certain category correctly, which means it decreases the impurity degree. The MDG can be calculated between 0 and 100. If one of the variables' MDG is 0, that variable is not used for classification at all. However, if the MDG is closer to 100, the variable can completely classify the observation.

5. Results

In the first stage, the MDG values of each input variable were calculated by setting the driver's type as a target variable and setting 37 characteristic variables of taxi drivers as input variables. The top 15 variables with large MDG values were extracted, and these variables were used as input variables for the next stage analysis. In the second stage, extracted variables from the first stage are set as input variables, and the accident risk group was set as a target variable.

5.1. Identification of factors characterizing taxi driver type

For learning the random forest, the total number of trees was set to 500, and each tree was set to build a tree using five variables at random. The OOB error of the model was 17.74%. Using the MDG values of each variable, the importance of each variable was assessed. Based on the MDG values, the top 15 variables considered to be effective in classifying taxi driver types were selected, and the results are presented in Table 1. The selected variables included in terms of health, work, and living environment of taxi drivers, such as the degree of mental health, the satisfaction of work, the stress level at work, fatigue, the satisfaction level for life, or the satisfaction level for the wage system.

Except for age, taxi driving experience, weekly working hours, and average number of weekend work per month, which can be represented by continuous variables, other variables were survey by 5-point scale. 5-point satisfaction obtained from mental health, job satisfaction, life satisfaction, health care, wage satisfaction, sleeping time, and vehicle interior environment indicates the best condition. On the other hand, the worst condition is represented by 5-points in case of job stress, fatigue, physical burden, and stress from passengers. Among these variables, the MDG value of the degree of mental health of the taxi drivers was analyzed to be 54.49 and was selected as the most effective factor in classifying the taxi driver's type. The MDG values of the satisfaction of work, stress level at work, fatigue, and the degree of the physical burden of taxi driving were calculated to be over 20, so it was found that the variables related to work and living environment are effective in classifying taxi driver type.

5.2. Prioritizing factors influencing accident risk

To analyze the factors influencing the classification of the accident risk group, a random forest methodology was applied using the 15 variables derived above as explanatory variables. The total number of trees was set to 500, and each tree was set to build a tree using two variables at random. The OOB error of the model was 30.75%. The results of the MDG values are presented in Fig. 4.

It was determined that the taxi driver's age has the most considerable influence on the classification of accident risk groups, with an MDG value of 28.4. In addition, the MDG values of the taxi driving experience of commercial vehicles, the satisfaction level for life, and the satisfaction of work were calculated to be over 20, so these variables were highly related to the classification of accident risk group. As shown in Fig. 4, the age and career of the taxi driver and the living and working environment were considered to have a significant influence on the taxi driver's accident risk group classification. We were able to extract 15 variables that can clearly

Table 1
Results of MDG by factors affecting taxi driver characterization.

Rank	Variable	Mean Decrease Gini
1	Degree of Mental Health	54.49
2	Satisfaction of Work	31.75
3	Stress Level at Work	27.21
4	Fatigue	25.9
5	Physical Workload	24.1
6	Satisfaction Level for Life	19.75
7	Health Management Level	18.47
8	Satisfaction Level for Wage System	16.38
9	Mental Stress due to Passengers	16.32
10	Age	15.72
11	Taxi driving Experience	13.63
12	Weekly Working Hours	11.68
13	Daily Sleeping Time	10.7
14	Monthly Weekend Work days	10.1
15	Satisfaction with Vehicle Interior Environment	8.78

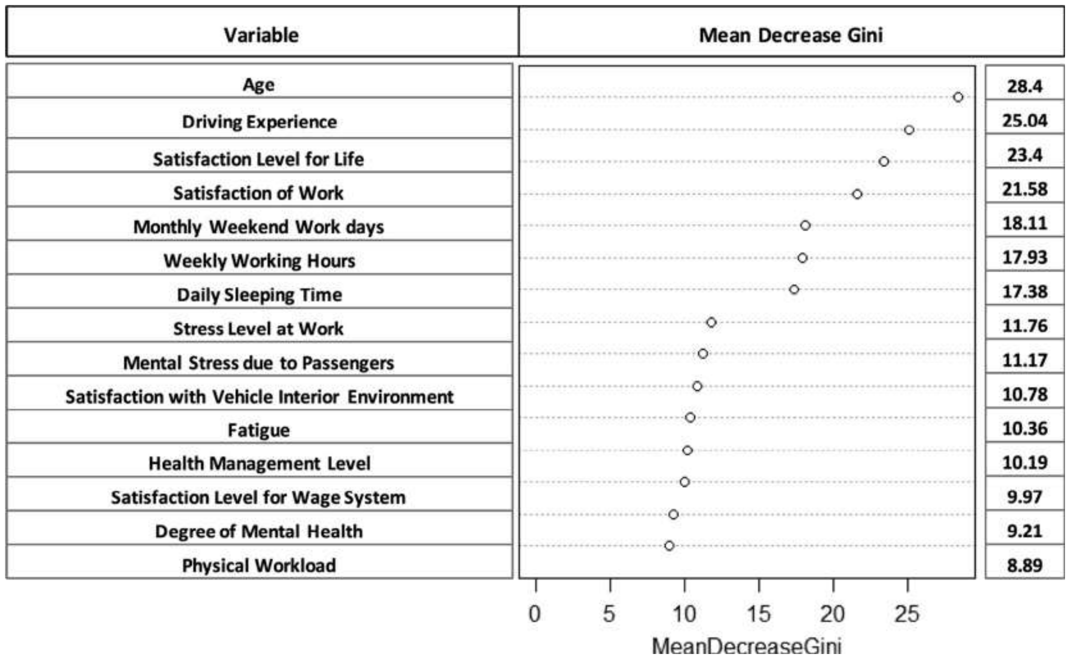


Fig. 4. Result of prioritizing factors to identify accident risks.

distinguish the difference among four types of taxi drivers through conducting the first random forest analysis. The result implied that mental health, job satisfaction level, and job stress contributed greatly to explaining the difference between driver types. The second random forest analysis was conducted to identify how much each variable would affect the crash severity. It was found that age, driving experience, and life satisfaction level had a great influence on the crash risk of taxi drivers. This study attempted to extract variables that can clearly classify driver types and driver's crash risk through the proposed methodology based on 2-stage random forest analysis.

5.3. Policy implications for improving taxi drivers' safety

Fifteen variables were selected as the meaningful variables, which were highly prioritized based on the RF analysis in this study. For example, the age and taxi driving experience of taxi drivers were categorized into demographic variables to represent the personal characteristics of drivers. Monthly weekend workdays, weekly working hours, and the satisfaction level for the wage system were regarded as variables for the work environment. Moreover, variables including daily sleeping time and the satisfaction level for life were interpreted as variables representing the health and welfare characteristics of taxi drivers. In this context, the 15 variables can be categorized into three groups: the taxi driver's personal characteristics, work environment characteristics, and health and welfare characteristics. This study derived a set of policy implications for the prevention of traffic accidents based on the analysis of variables influencing the classification of taxi driver types in terms of traffic safety. Details on each policy implication are presented in the subsequent subsections. The relationship between driver's intrinsic factors and policy implications is presented in Fig. 5.

5.3.1. Establishing a driving capability evaluation system for elderly drivers

A systematic evaluation of the driving capability of elderly drivers is required to identify high-risk taxi drivers. With the evaluation of driving ability, continuous monitoring of work and life environments and health conditions are also useful in developing effective countermeasures to enhance traffic safety associated taxi drivers. As a part of the evaluation system, it is recommended to add evaluation items to represent the work and health-related conditions.

5.3.2. Improving the driver's wage system

It is necessary to improve the satisfaction of the driver's wages through a policy that guarantees the minimum wage of taxi drivers, which leads to the reduction in work-related stress. This ultimately contributes to enhancing traffic safety.

5.3.3. Restricting overtime work for taxi drivers

One of the important factors resulting in accidents is the fatigue of drivers. A policy restriction on overtime work for taxi drivers is necessary to significantly reduce fatigue, including physical and mental stresses.

5.3.4. Supporting the improvement of rest facility

Providing shared rest facilities for taxi drivers is necessary to improve the work environment toward relieving the fatigue. In the case of a small taxi company, the remaining facilities are often in poor condition. A feasible solution to this problem is to provide the shared rest facility at the local government level.

5.3.5. Developing mental health management programs for taxi drivers

The purpose of this policy is to reduce the stress level at work and the crash risk by establishing a system to monitor and manage the mental health of taxi drivers. Through collaboration between local public health centers and taxi companies, it is necessary to periodically monitor the physical and mental health condition of taxi drivers. In addition, it is expected that the quality of life and traffic safety of taxi drivers will be improved by providing guidelines for mental health management.

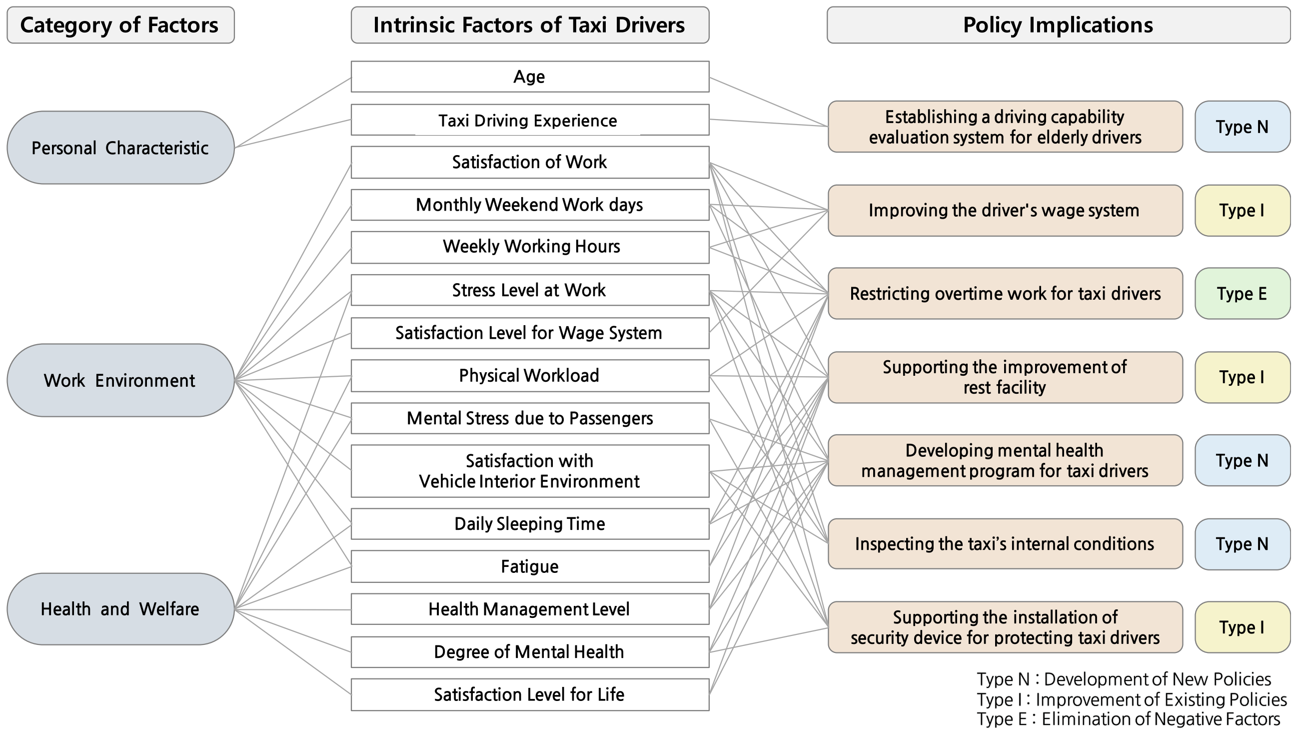


Fig. 5. Mapping driver's intrinsic factors with policy implications.

5.3.6. Inspecting the taxi's internal conditions

This policy is to improve the in-vehicle condition. Specifically, it is necessary to establish an evaluation system to maintain the cleanliness of taxis at the governmental level to perform the periodic evaluation. In addition, it is necessary to check the condition of the inside of the vehicle together with regular vehicle inspection.

5.3.7. Supporting the installation of security devices for protecting taxi drivers

This policy ensures the safety of taxi drivers who are exposed to physical and mental damage from passengers due to the nature of the work. It is possible to ensure the safety of taxi drivers by establishing a regulation that legally mandates the installation of CCTV in the taxi. To reduce the economic burden of taxi drivers as much as possible, the subsidy for the installation of the protection device should be increased.

The seven policies for taxi drivers derived in this study can be classified into three types: N, I, and E according to the characteristics of the policy. Type N is a new policy developed for taxi drivers, and Type I is a policy that promotes improvements of the existing policy for taxi drivers. Lastly, Type E is a policy to eliminate negative factors for the safety of taxi drivers. Each of the three types of policies has a direction of policy promotion based on the current situation and problems.

Currently, efforts for developing and implementing policies for evaluating the driving ability of taxi drivers and managing the working environment in Korea are insufficient. Institutional measures to improve the driving ability and working environment are needed because they are directly related to crash risks. In this study, a new policy that can be prepared for taxi driver work environment and traffic safety management is presented as Type N. One of the critical issues for taxi drivers who drive long hours in limited places and deal with various customers is security and rest. Type I is a policy that can improve traffic safety for taxi drivers by improving existing policies. For example, efforts are needed to enhance the effectiveness of policies to ensure rest and prevent violence from passengers. In addition, it is necessary to improve the wage system because work stress and overworking leading to crash risk can be reduced by devising proper wage system. Finally, Type E refers to a policy to reduce exposure to traffic crashes by removing the factors that hinder traffic safety of taxi drivers. It is believed that with systematic restrictions of overworking taxi drivers, it will be possible to not only reduce crash risks but also improve the quality of life.

In this study, we analyzed the priority of the proposed policy implications to emphasize the need for institutional management for taxi drivers. To prioritize the seven policies for improving traffic safety for taxi drivers, the results of the priority analysis of the factors affecting the classification of accident risk groups were used. First, a process of matching explanatory variables closely related to each policy was performed. For example, the variables of the age of driver and taxi driving experience were matched with the policy for the establishment of an evaluation system for the driving capability of the elderly taxi driver. The variables for the satisfaction of work, the average number of working days on the weekend per month, the stress level at work, and the satisfaction level for the wage system were matched with the policy for the improvement of the taxi driver's wage system.

After analyzing the relationship between 7 policies and 15 explanatory variables, the priority was derived by calculating the average value of MDG for each policy. The equation for the calculation of the average MDG of each policy is presented in Equation 3.

$$\overline{MDG_i} = \frac{\sum_{k=1}^{n} MDG_a}{N}$$

i indicates policy and $\overline{MDG_i}$ indicates the average MDG of policy i. MDG_a indicates the MDG value for each explanatory variable, and N is the number of explanatory variables matched with the policy. The results of the average MDG for each policy are presented in Table 2, and as a result of the analysis, it is analyzed that the

Table 2
Results of average MDGs by policy implications.

Rank	Policy Implications	$\overline{MDG}$
1	Improving the driving capability evaluation system for elderly drivers	26.72
2	Improving the driver's wage system	16.90
3	Restricting overtime work for taxi drivers	14.53
4	Improving driver's rest environment	14.10
5	Developing mental health management program for taxi drivers	14.00
6	Improving the taxi's internal conditions	13.25
7	Supporting the installation of security device for protecting taxi drivers	12.90

establishment of an evaluation system for the driving ability of the elderly taxi driver is preferentially necessary to improve the traffic safety of the taxi driver.

6. Conclusion

Traffic safety issues associated with taxis are of keen interest because the frequency of taxi crashes is significantly higher than that of other vehicle types. Thorough investigations on the cause of taxi crashes are fundamental to developing effective policy-based countermeasures. A variety of existing studies have attempted to identify influencing factors leading to crashes. Variables related to drivers' human factors have been known to be dominant and critical among such factors. However, most existing studies deal with external causal factors indicative of unsafe driving behavior characteristics resulting in crashes, such as inattention, distraction, drowsiness, and aggressiveness. Unlike existing studies, this study focuses on underlying intrinsic aspects that can be the root cause of unsafe driving behavior. In addition to work and living environment, drivers' health-related conditions, obtained from an in-depth questionnaire survey, were analyzed to identify intrinsic factors contributing to taxi safety. It should also be noted that this study proposed a methodology based on an RF technique, which is a widely used heuristic method without inevitable assumptions for statistical analyses, to mine meaningful implications to be used for deriving safety policies.

The methodology proposed in this study extracted risk factors by applying RF in two stages. In the first stage, factors capable of distinguishing the type of taxi driver were identified. It was found that a total of 15 variables were capable of characterizing taxi driver types. Next, extracted variables from the first stage were used as inputs, and the crash risk level including high-risk, medium-risk, and low-risk was set as a target variable to prioritize variables in terms of the importance level affecting the safety of taxi drivers. This study derived a set of policy implications for the prevention of taxi crashes from three perspectives: personal, working, and health and welfare. In addition, the main purpose and key issues of each policy were discussed. The derived policies can be divided into the development of new policies, the improvement of existing policies, and the elimination of negative factors. Regarding the development of new policies, this study proposed 'establishing a driving capability evaluation system for elderly drivers,' 'developing mental health management programs for taxi drivers,' and 'supporting the installation of security device for protecting taxi drivers.' Additionally, 'improving the driver's wage system,' 'inspecting the taxi's internal conditions,' and 'supporting the improvement of rest facility' are for improving existing policies. Last, the main purpose of 'restricting overtime work for taxi drivers' is to eliminate negative factors. It is believed that the outcome of this study will be valuable in developing policy countermeasures to prevent taxi crashes. However, further studies need to be conducted to obtain more substantial and useful implications for safety policies. First, more sample data should be collected with the institutionalization of the in-depth survey. It is also necessary to continuously discover new questions and improve existing questions to be used for identifying intrinsic factors associated with taxi safety. From the methodological point of view, performing the proposed methodology with various different types of target variables is expected to allow for obtaining more meaningful results. In particular, different definitions need to be applied to classify the risk level of taxi drivers. In addition, among all licensed taxis in Korea, corporate taxis and private taxis account for 34.53% (86,935 taxis) and 65.47% (164,868 taxis), respectively (2020, National Association of Taxi Transportation Business Associations). It is expected that there exist obvious differences between corporate taxis and private taxis in various aspects such as criteria for obtaining a license, wage system, and working hours, which will affect the characteristics related to health, work, and living conditions of taxi drivers. Therefore, it is necessary to expand our study subject to private taxis in the future, as this study only focused on corporate taxis. This study conducted a 2-stage random forest analysis, which was focused on the identification of the relationship between the health, work, and life characteristics and traffic safety of individual corporate taxi drivers. While conducting this analysis, the geographical characteristics of taxi companies and drivers were not considered. However, it is expected that the geographic environment would affect the personality and attitude contributing to driving behavior. Therefore, a future study should take geographic environment into consideration in analyzing taxi crashes and drivers' internal characteristics.

Acknowledgement

This research was supported by a grant from Transportation and Logistics Research Program funded by Ministry of Land, Infrastructure and Transport of Korea government (20TLRP-B148659-03).

References

Abdel-Aty, M., Pande, A., Das, A., & Knibbe, W. J. (2008). Assessing safety on Dutch freeways with data from infrastructure-based intelligent transportation systems. *Transportation Research Record, 2083*(1), 153–161.

Al-Ghamdi, A. S. (2000). Taxi service characteristics and involvement in road traffic accidents in Riyadh. *Journal of King Saud University-Engineering Sciences, 12*(2), 199–216.

Apantaku-Onayemi, F., Baldyga, W., Amuwo, S., Adefuye, A., Mason, T., Mitchell, R., & Blumenthal, D. S. (2012). Driving to better health: Cancer and cardiovascular risk assessment among taxi cab operators in Chicago. *Journal of Health Care for the Poor and Underserved, 23*(2), 768.

Borowsky, A., & Oron-Gilad, T. (2013). Exploring the effects of driving experience on hazard awareness and risk perception via real-time hazard identification, hazard classification, and rating tasks. *Accident Analysis & Prevention, 59*, 548–565.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

Bulduk, E. Ö., Bulduk, S., Süren, T., & Ovalı, F. (2014). Assessing exposure to risk factors for work-related musculoskeletal disorders using Quick Exposure Check (QEC) in taxi drivers. *International Journal of Industrial Ergonomics, 44*(6), 817–820.

Chen, J. C., Chen, Y. J., Chang, W. P., & Christiani, D. C. (2005). Long driving time is associated with haematological markers of increased cardiovascular risk in taxi drivers. *Occupational and Environmental Medicine, 62*(12), 890–894.

Deng, H., Runger, G., & Tuv, E. (2012). System Monitoring with Real Time Contrasts. *Journal of Quality Technology, 44*(1), 9–27.

Do, C. W. (2013) Transportation Safety Engineering Cheongmoongak, pp.105–111.

Hansen, J., Raaschou-Nielsen, O., & Olsen, J. H. (1998). Increased risk of lung cancer among different types of professional drivers in Denmark. *Occupational and Environmental Medicine, 55*(2), 115–118.

Harb, R., Yan, X., Radwan, E., & Su, X. (2009). Exploring precrash maneuvers using classification trees and random forests. *Accident Analysis & Prevention, 41*(1), 98–107.

Jiang, X., Abdel-Aty, M., Hu, J., & Lee, J. (2016). Investigating macro-level hotzone identification and variable importances.

Lee, M. W., Kim, Y. G., Jun, Y. J., & Shin, Y. H. (2019). Random Forest based Prediction of Road Surface Condition Using Spatio-Temporal Features. *Journal of Korean Society of Transportation, 37*(4), 338–349.

Lee, S. Y., Cho, E. S., Oh, C., Kang, S. H., & Choi, B. H. (2020). Pattern Classification of Taxi Drivers Based on Work and Living Conditions for Commercial Vehicle Safety Management System. *Journal of Korean Society of Transportation, 38*(3), 218–231.

Li, M. K., Yu, J. J., Ma, L., & Zhang, W. (2019). Modeling and mitigating fatigue-related accident risk of taxi drivers. *Accident Analysis & Prevention, 123*, 79–87.

Lim, S. M., & Chia, S. E. (2015). The prevalence of fatigue and associated health and safety risk factors among taxi drivers in Singapore. *Singapore Medical Journal, 56*(2), 92.

Maag, U., Vanasse, C., Dionne, G., & Laberge-Nadeau, C. (1997). Taxi drivers' accidents: How binocular vision problems are related to their rate and severity in terms of the number of victims. *Accident Analysis & Prevention, 29*(2), 217–224.

Ma, J., & Cheng, J. C. (2016). Identifying the influential features on the regional energy use intensity of residential buildings based on Random Forests. *Applied Energy, 183*, 193–201.

Meng, F., Wong, S. C., Yan, W., Li, Y. C., & Yang, L. (2019). Temporal patterns of driving fatigue and driving performance among male taxi drivers in Hong Kong: A driving simulator approach. *Accident Analysis & Prevention, 125*, 7–13.

Miyamoto, M., Konno, S., Gembun, Y., Liu, X., Minami, K., & Ito, H. (2008). Epidemiological study of low back pain and occupational risk factors among taxi drivers. *Industrial Health, 46*(2), 112–117.

Siddiqui, C., Abdel-Aty, M., & Huang, H. (2012). Aggregate nonparametric safety analysis of traffic zones. *Accident Analysis & Prevention, 45*, 317–325.

Vahedi, J., Shariat Mohaymany, A., Tabibi, Z., & Mehdizadeh, M. (2018). Aberrant driving behaviour, risk involvement, and their related factors among taxi drivers. *International Journal of Environmental Research and Public Health, 15*(8), 1626.

Jiwon Park received the B.S. degree from Hanyang University in 2018, where she is currently working toward the Ph.D. She is currently studying traffic safety and crash data analyses.

Seolyoung Lee received the Ph.D. degree from Hanyang University in 2020. She is currently a principal researcher of the department of Smart City Research at Seoul Institute of Technology, Korea. Her research interests include traffic crash data analysis, traffic and driving simulations, and traffic safety issues in mixed traffic flow with automated vehicles and regular vehicles.

Cheol Oh received the Ph.D. degree in civil engineering–transportation systems engineering from the University of California at Irvine in December, 2003. He is currently a full professor of the department of transportation and logistics engineering at Hanyang University, Korea. He is conducting research on commercial vehicle safety, active traffic safety analytics, and traffic management strategy under cooperative automated driving environments.

Byongho Choe received the Ph.D. degree in psychology from the Oldenburg University in 2001. He is currently a director of the transportation safety research and development institute of Korea Transport Safety Authority. His research includes traffic safety policies, driving behavior analysis, and crash data analysis and management systems.

Journal of Safety Research 76 (2021) 166–175

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Adolescent noncompliance with age-specific versus universal US motorcycle helmet laws: Systematic review and meta-analysis

Jon Berrick [a,*], Konstantina Gkritza [b]

[a] Science Division, Yale-NUS College, 10 College Avenue West #01-101, Singapore 138609, Singapore
[b] Lyles School of Civil Engineering and Department of Agricultural & Biological Engineering, Purdue University, 550 Stadium Mall Dr, West Lafayette, IN 47907-2093, USA

ARTICLE INFO

Article history:
Received 23 April 2019
Received in revised form 18 September 2020
Accepted 15 December 2020
Available online 2 January 2021

Keywords:
Adolescent psychology
Age-restricted
Attributable percentage among exposed
Threat of apprehension
Youth control

ABSTRACT

Introduction: The U.S. experience with motorcycle helmets affords an important insight into the responses of adolescents to age-specific laws. Political contention has led to a number of U.S. state law changes back and forth between universal and age-specific laws. Because both kinds of law require adolescent motorcyclists to wear helmets, relatively few studies have focused on how the law type affects their behavior. *Method:* Differential behavior is tested by a systematic review of literature, leading to a meta-analysis, in relation to the experience of various states' motorcycle helmet laws. An electronic search was conducted for before-and-after studies in U.S. states that include data on adolescent helmet usage – both with a universally applicable motorcycle helmet law, and with an age-restricted law (usually, under-21 or under-18) – from observational, injury or fatality records for a certain period (e.g., 12 months) pre and post the state law change. *Results:* The search yielded ten studies, including two that compared a set of age-specific law states with a set of universal law states over the same time period. Heterogeneity analysis of seven single-state studies with raw data revealed an acceptable fit for a random-effects model. Additional noncompliance with age-restricted laws was indicated by an attributable percentage among exposed of over 65% and odds ratio exceeding 4. *Conclusions:* About two-thirds of adolescent noncompliance with age-restricted motorcycle helmet usage laws disappears with universal applicability. Evidence from numerous international studies of youth reaction to helmet laws suggests that a large part of the greater compliance with universal laws is due to their conveying a more convincing message that helmets afford protection against injury. *Practical Applications:* The meta-analysis provides fresh, young-rider perspective on the continuing debate over motorcycle-helmet laws. Broader insight into adolescent psychology suggests considering alternatives to age-restricted laws more widely in safety and health policy.

1. Introduction

There are a number of public health and safety domains where the behavior of adolescents has an important role. Often the regulatory response to this lies in age-specific laws. But how effective are they? In many cases (e.g., alcohol and tobacco), it is difficult to draw direct conclusions because there are few instances of contrasting laws within comparable populations. United States motorcycle helmet laws are an important exception: many states have had different regimes in force at different periods, offering an opportunity for making comparisons.

1.1. Motorcycle helmets

Motorcycles are a distinctly hazardous form of travel. By distance traveled, U.S. motorcyclists are over 30 times more likely to die in a crash than drivers of other types of motor vehicles (Lin & Kraus, 2008); in 2018, 3% of motorcycle fatalities involved riders 19 years or less (IIHS, 2019). The role of helmets in preventing motorcyclist fatalities and alleviating injury has long been demonstrated (Evans & Frick, 1988). An international systematic review (Liu et al., 2008) concluded that helmets reduce the risk of head injury by 69% and death by 42%. In turn, universal helmet laws are recognized as one of the most effective approaches for increasing helmet use (Lee, 2018; Peng et al., 2017) and their benefits persist over the long term (Lee & Outlaw, 2018). Accordingly, numerous countries have adopted such policies, including those of the European Union, Japan, Canada, Australia, Singapore, and New Zealand.

* Corresponding author at: School of Mathematics and Statistics, F07, University of Sydney, NSW 2006, Australia. School of Computer, Data and Mathematical Sciences, Western Sydney University, Locked Bay 1797, Penrith NSW 2751, Australia.

E-mail address: jon.berrick@yale-nus.edu.sg (J. Berrick).

https://doi.org/10.1016/j.jsr.2020.12.011
0022-4375/

1.2. Importance of adolescent compliance

It is well known that adolescents are overrepresented in motor-vehicle crashes (Shope & Bingham, 2008). Obviously less experienced, they are also considered to be more liable to sensation-seeking and impatience (Wong et al., 2010). They could further be described as less risk-aware: a West Australian survey indicated that while 85% of the general community expressed concern about the safety of novice drivers, only a minority of novice drivers shared this concern (OECD/ECMT, 2006). Lowered risk perception by adolescents has been attributed to personality and social norms (Falco et al., 2013), sleep deprivation (Groeger, 2006), and socio-cultural contributors (Njå & Nesvåg, 2007).

Evidently then, to the extent that helmets are protective, it is important that youth obey laws mandating their use. In the Discussion section, we seek to analyze various contributors to noncompliance, informed by broad theoretical models of motivators of behavior. It is clearly valuable to gain an understanding of factors affecting compliance that fall under the control of policy-makers. Is it possible that the nature of the law itself might influence its acceptance? An affirmative answer could have implications beyond the immediate realm of adolescent motorcycle safety.

1.3. Universal versus age-specific laws for increasing adolescent motorcycle helmet use

The decade following the U.S. Highway Safety Act of 1966 saw almost all U.S. states respond to federal economic incentives to introduce universally applicable motorcycle helmet laws. Following the discontinuance of this support with the 1976 Federal-Aid Highway Act, and under pressure from motorcycle rights organizations, over succeeding decades the majority of states restricted the applicability of their laws to adolescents only (typically those under 18 or under 21; Lee & Outlaw, 2018). Continued political contention has led to a number of law changes back and forth between universal and age-specific laws (Dee, 2009; Homer & French, 2009; Vaca, 2006).

Adoption of partial laws has resulted in fewer riders wearing helmets, and increased motorcyclist fatalities and injuries experienced by those states that repealed universal helmet laws (Buckley et al., 2016). In states with an under-21 law, Weiss et al. (2010) found that serious traumatic brain injury among young riders was 38% higher than in states with universal helmet laws.

1.4. Existing literature and previous related reviews/meta-analyses

Political debate about appropriate helmet laws has largely contrasted issues of effectiveness of helmets in reducing injury with libertarian concerns (Jones & Bayer, 2007). Consequently, there are numerous studies of effects of such law changes. However, instead of adolescent response they tend to focus on questions relating to number of lives lost (Lee et al., 2017), frequency (Olsen et al., 2016) and nature (Peek-Asa & Kraus, 1997) of injuries (Abbas et al., 2012), and economic costs (Kim et al., 2015); see also (Coben et al., 2007).

Because adolescent motorcyclists are required to wear helmets under both kinds of law, some authors have assumed that relevant statistics would not vary according to the type of law. For example:

- "Since the change in helmet wearing requirements occurred for individuals 18 years of age and older, information was examined for those individuals only" (Fleming & Becker, 1992, p.835);
- "Because Louisiana has always mandated helmet use in the underage group, pre- and post-repeal helmet use in the under-18 age group would be expected to be unchanged" (Ho & Haydel, 2004, p.154);
- "Of course, because the repeal of the motorcycle helmet-use law applied only to motorcycle operators and passengers older than 21 years of age, the repeal of the law should have little if any effect on the serious injury and fatality rate series for motorcycle riders younger than 21 years of age. The use of the younger-than-21 serious injury and fatality rates as statistical controls in the analyses helps us to avoid attributing significance to the repeal of the law that should, more accurately, be attributed to some other independent but coincidental event such as a change in weather conditions" (Stolzenberg & D'Alessio, 2003, p.134).
- On the other hand, Watson et al. (1980, p.581) ignores underage data "because the number of motorcyclist deaths per month in each age class would have been too small."

Accordingly, relatively few studies appear to have attempted to measure the effect of law changes on youth helmet usage, or even to have published enough data to allow the effect to be measured.

Two multi-state reviews that have considered youth accordance with different laws are Brooks et al. (2010) and Peng et al. (2017). The former analyzed motorcycle fatality data from the years 1996–2005; for those under 18 years, the study reported no significant difference in fatality rates per 10,000 registrations between the three states with no helmet laws and three states with partial helmet laws (states with universal helmet laws were not considered). The latter is a systematic review (without meta-analysis), which noted sizeable percentage point differences in youth helmet usage between generic and age-limited regimes.

1.5. Research objectives

The bulk of the present research comprises a systematic review of literature, leading to a meta-analysis, to address the following question (Berrick PROSPERO, 2018):

In relation to the experience of various US states' motorcycle helmet laws, what is the effect on adolescent noncompliance of an age-specific law instead of a universal law?

The approach is to consider cohort studies that include data on adolescent helmet usage for a certain period (e.g., 12 months) both pre- and post- a particular state law change. These may be ascertained from observational, injury or fatality records.

Secondly, insofar as the meta-analysis reveals a difference in outcomes, we then seek to discuss what factors seem to be relevant in leading to such a difference. The observation of Houston (2007, p.334) that "partial coverage statutes undermine the motives that lead individuals to comply with the law," indicates the need for examination of the motives for adolescent noncompliance with age-specific laws. Although this discussion is necessarily more tentative than the preceding meta-analysis, we feel that such insights can be of value to the wider public health and safety community in other domains where age-specific laws prevail. To frame the discussion, we now briefly review the theoretical background.

1.6. Helmet laws and adolescent psychology: four hypotheses

The first two common surmises are those that accord with deterrence theory, as discussed in Houston (2007), whereby it is the likelihood of punishment that most leads to compliance.

(i) Conspicuity. Generic laws make all unhelmeted riders subject to police apprehension, increasing the deterrent impact in comparison with an age-specific regime where it might not be easy for police to decide the age of an unhelmeted rider.

(ii) Enforcement prioritization. "Less rigorous enforcement may also result from perceived lack of priority once older age groups have been exempted from helmet-use compliance" (Weiss et al, 2010, p.1593).

In contrast, the theory of planned behavior (Ajzen, 1991) and social learning theory and the prototype/willingness model (Scott-Parker et al., 2013) highlight the influence of subjective and social norms in giving laws their expressive effects. See also Fig. 1 "Typology of determinants of adherence to a law" in Berrick (2013).

(iii) Mixed messaging. The "perceived lack of priority" above might also influence the motorcyclists themselves; in other words, they may interpret an age-specific law as a signal that lawmakers are less convinced of the dangers of riding unhelmeted.

Lastly, reactance theory (Dillard & Shen, 2005) suggests that the discriminatory nature of partial laws hints that their ulterior purpose is youth control and may thus incite rebellion among those who feel discriminated against.

(iv) Rite-of-passage. The requirement to wear a symbol of one's immaturity contradicts the expression of independence signified by the mobility of ridership.

Evidently, the above issues are relevant to any choice between a generic and an under-age law.

2. Materials and methods

An electronic search was conducted for studies (1970s until 2018) in U.S. states that have changed their motorcycle helmet laws between an age-restricted and a universal law, in order to gather statistical data comparing adolescent noncompliance under each condition. This was followed by a meta-analysis of the data obtained.

2.1. Criteria for considering studies for this review

2.1.1. Types of studies

Studies sought were cohort studies that include data on adolescent helmet usage for a certain period (e.g., 12 months) both pre and post- the state law change. These may be ascertained from observational, injury, or fatality records.

2.1.2. Participants/population

In a U.S. state that has changed from one kind of law to another, the relevant population comprises motorcyclists under the age mandated to wear helmets in the age-restricted regime (usually, under-21 or under-18).

2.2. Comparator(s)/control

The search sought before-and-after studies publishing rates of adolescent noncompliance – both with a universally applicable motorcycle helmet law, and with an age-restricted law. (Due to the symmetry of the situation, either the universal or age-specific regime could be considered the control in the meta-analysis.)

2.3. Context

By studying data for individual U.S. states, interstate demographic differences are controlled for. As most studies relate to a year or two before and after the law change, this reduces the impact of temporal demographic changes.

2.4. Search methods for identification of studies

Two reviewers worked independently to obtain suitable studies from databases, relevant journals, previous reviews, and transportation agency websites. Further sources were obtained by (mostly backward) snowballing. Where appropriate, authors were contacted in order to obtain raw data, consisting of actual numbers of crash-involved persons.

2.4.1. Inclusion criteria

The search filtered for the terms "motorcycle" and "helmet," but was further refined by terms such as "law" and "compliance" in the case of the databases MEDLINE, Google Scholar, The Cochrane Library and nhtsa.gov., as well as the U.S. General Accounting Office. Journals searched were Accident Analysis & Prevention, American Journal of Public Health, Annals of Emergency Medicine, Injury Prevention, Journal of Health Economics, Journal of Safety Research, Journal of Trauma–Acute Care Surgery, Journal of Trauma–Injury Infection & Critical Care, The American Journal of Surgery, and Traffic Injury Prevention.

2.4.2. Exclusion criteria

From studies initially retrieved there were excluded those focusing on such matters as all-terrain vehicles, bicycles, seatbelts, pedestrians, young children, helmet design, alcohol, nature of injury, economic impacts, education, and non-U.S. data. Studies with very small samples were excluded; and in the case of more than one study covering a particular law change, so as to avoid double-counting, validity and sample size were considered in order to determine which one study would be included.

2.5. Assessment of risk of bias in included studies

When the question of relative adolescent noncompliance is not the primary objective of a study, its results of interest to this review tend to be reported only incidentally. This reduces the likelihood of *publication bias*. The main reasons why the question is unlikely to be the primary objective of a study are considered to be the following.

(i) It is counter-intuitive: when adolescents are subject to both age-specific and universal laws, theoretically their compliance should not vary.
(ii) Most studies are primarily interested in direct safety and economic impacts of helmet laws, rather than compliance and its causes.
(iii) Adolescents form a small portion of the motorcyclist population. Since most studies focus on injuries or economic consequences, there is little motivation to isolate this relatively small subgroup.
(iv) Given the unreliability of observational assessment of age of motorcyclist, a much larger overall sample is needed in order to obtain usable data for the adolescent subgroup.

Selection bias is not a major issue because the whole adolescent population is subject to both laws. However, in a study spanning several years, the adolescent population at the end of the period will not comprise the same people as those at the beginning of the study.

Attrition bias could arise because an unhelmeted cyclist is more likely to suffer a fatal injury, suggesting that, ceteris paribus, the proportion of cyclists using helmets would rise over time. However, given the low frequency of fatalities relative to demographic movements, any such effect is not considered likely to be significant.

There is some evidence that observation data and crash data yield comparable predictions of helmet use. For example (Buckley et al., 2016): "The current study estimated the rate of helmet use (including novelty helmets) at 75% for all motorcycle operators, compared with 73% identified in police-reported crash data.

The similarity and alignment is notable and suggests there is value in using crash data, in particular data that include property-damage-only crashes, when understanding statewide helmet use rates. Given the costs of large scale observations and that crash data are routinely collected by states, helmet use rates in crash data may be a useful and inexpensive way to approximate overall helmet use and especially to identify associated characteristics and changes over time."

Detection bias can be an issue when injury data are recorded, if records rely on patients' self-reporting, because helmet usage is mandatory. However, this applies equally to adolescents under both helmet laws. The bias can be reduced by relying on fatality data instead, when available.

The *statistical bias* created by the protective effect of helmets in using fatality data to determine rates of helmet usage is addressed in the Appendix.

A bias due to an underlying *trend* of helmet use is mitigated by studying both directions of change – from age-limited to universal and from universal to age-specific laws. Moreover, we consider national studies that examine over a uniform time period adherence to each kind of law in a variety of states.

Performance bias is a possibility when comparing records several years apart, because it may be that the records were taken according to different protocols. Nevertheless, the clarity of the dichotomous outcome – using or not using a helmet – mitigates this. There may be changes over time in recorders' interpretations of what constitutes a "proper" helmet; this is not considered to be a significant factor.

2.6. Effect measures

The statistics used in this analysis are attributable percentage among exposed (APaE) and odds ratio (OR). Their suitability is discussed as follows.

A typical policy-maker's question is: By what percentage will a current adverse outcome (noncompliance in the present instance) be reduced if the proposed intervention is introduced? (Compare Rockhill et al., 1998; Gefeller, 2001.) It is addressed here by the statistic known as attributable percentage (or proportion, or fraction, or risk) among exposed (Cole & MacMahon, 1971). It is probably most readily understood as a percentage, which we refer to as APaE. Its formula is

$$\mathrm{APaE} = 100 \times (R_c - R_p)/R_c = 100 \times (1 - 1/RR) \quad (1)$$

Here, R_c and R_p denote respectively the noncompliance rates under current and proposed policies, while RR is the relative risk R_c/R_p.

While the odds ratio is considered to have relatively attractive theoretical properties, it is less readily explained to a policymaker. A possible explanation might be as follows. For each policy, consider the multiple by which compliance outweighs noncompliance; then record the relative size of those two multiples. The formula is

$$\mathrm{OR} = R_c(1 - R_p)/(R_p(1 - R_c)) \quad (2)$$

2.6.1. Illustrative example

As a hypothetical example, if the proposed intervention was associated with a lowering of the noncompliance rate from its current 25% to 10%, that would count as a reduction of 60% (being 15/25), so APaE = 60%. In this example, the current policy has compliance three times as large (being 75/25) as noncompliance; for the proposed policy this multiple would be 9 times (being 90/10). Therefore, the odds ratio is 3 (being 9/3).

2.7. Strategy for data synthesis

Attributable percentage among exposed and odds ratio for noncompliance attributable to an age-restricted law (in contrast to a universally applicable law) were determined for each studied law change. This was a point estimate in the case where raw data were not available, and a 95% confidence interval where data were available. Comparison of studies led to a test for heterogeneity and of the suitability of a fixed-effect model (Borenstein et al., 2009). Subgroup analysis was to be conducted if considered appropriate from the results of the heterogeneity analysis. Using either a fixed-effect or random effects model (as found appropriate), for those studies where raw data exist to allow pooling of data, aggregated attributable risk and odds ratio were determined, together with 95% confidence intervals, displayed by forest plot.

Pooling was conducted only for (fixed state, varied regime) studies pertaining to pre- and post-law change data within a single state. Any (fixed regime, varied state) studies comparing a set of universal-law states with a set of partial-law states were excluded from heterogeneity analysis and pooling. This is because of the different nature of such studies: it may be, for example, that cultural differences between states that have a universal law throughout the study period and those that have an age-restricted law throughout the period act as a confounder.

3. Results

3.1. Search yield

Progress of the search for relevant studies is indicated in Fig. 1. Altogether, preliminary screening by key words and phrases as described in Methods above yielded 1,288 papers for further review. Further screening by title and abstract, and removal of duplicates, resulted in 138 articles assessed on the basis of their full text. After further exclusion of repetitious reports, so as to avoid double-counting of a single law change event, the remaining 10 studies, detailed in Table 1, comprised the basis for meta-analysis. There were no disagreements between the authors relating to study eligibility. Data extraction details are available from the corresponding author.

Two of the 10 selected studies did not relate to a single state law change, but compared, over a period of years, compliance in a sample of states with age-restricted laws against compliance in a set of states with age-unrestricted laws. Of the remaining eight studies, one did not provide raw data. The seven single-state studies with raw data were then subjected to heterogeneity analysis.

3.1.1. Adjustment for fatality data

The literature suggests that the effectiveness *e* of helmets in saving lives in motorcycle crashes has improved over time as helmet design has become more effective (Lin & Kraus, 2008). Since the years for which fatality data were used centered around 1977 (Texas DPS, 1991), 2000 (Ulmer & Northrop, 2005) and 1995–2003 (Mayrose, 2008) – see Table 1, the value of *e* indicated by NHTSA (1998), namely *e* = 36%, was used for this purpose. See the Appendix for details of how the value of *e* is applied.

3.1.2. Description of studies selected

Table 1 presents a summary of the studies included in the meta-analysis.

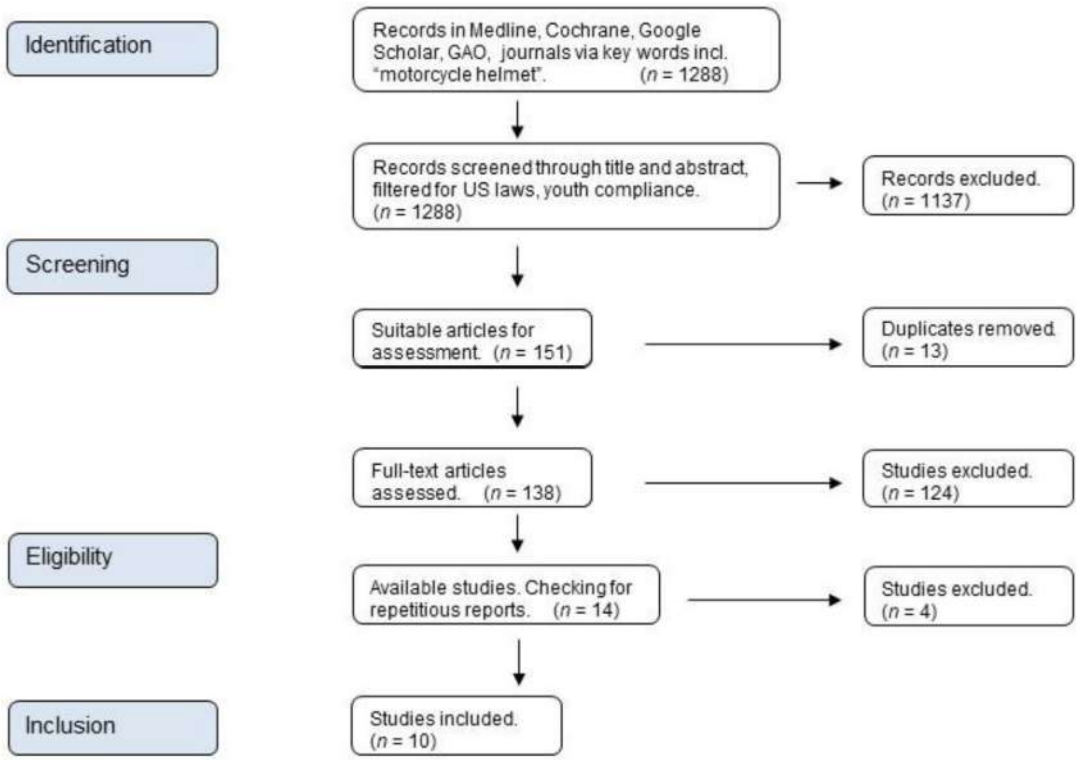


Fig. 1. Search results PRISMA flowchart.

Table 1
Studies included in the meta-analysis.

State and change year	Change	Data sources Pre-change	Data sources Post-change
SD77*	U → 18- July 1, 1977	Struckman-Johnson and Ellingstad, 1980: all crashes; months 12–0	Struckman-Johnson and Ellingstad, 1980: all crashes; months 0–24
ND77	U → 18- July 1, 1977	Heilman et al., 1982: injuries; months 6–0	Heilman et al., 1982: injuries; months 0–42
TX77	U → 18- August 29, 1977	Texas DPS, 1991: fatalities; months 12–0	Texas DPS, 1991: fatalities; months 0–12
LA82	18- → U January 1, 1982	McSwain and Belles, 1990: all crashes; months 12–0	McSwain and Belles, 1990: all crashes; months 0–12
TX89	18- → U Sept. 1, 1989	Goodnow, 1990: injuries; months 36–20	Preusser et al., 2000: injuries; months 52–64
TX97	U → 21- Sept. 1, 1997	Preusser et al., 2000: injuries; months 44–0	Preusser et al., 2000: injuries; months 0–16
FL00	U → 21- July 1, 2000	Ulmer and Northrop, 2005: fatalities; months 42–6	Ulmer and Northrop, 2005: fatalities; months 6–42
MI12	U → 21- April 12, 2012	Carter et al., 2017**: all crashes; months 27–3	Carter et al., 2017**: all crashes; months 9–33
USA95-03	Mayrose, 2008: fatalities		
USA05-08	Olsen et al., 2016: all crashes		

*: Insufficient data to enable calculation of 95% Confidence Intervals (thus excluded from pooling).
**: Supplementary data kindly provided by authors (enabling CI and pooling).
U: Universal law (no age-restriction).
18-: Law applies only to those under 18 years (possibly also to uninsured, etc.).

3.2. Meta-analysis

3.2.1. Study results

As described in (2.7) Strategy for data synthesis, attributable percentage among exposed (APaE) and odds ratio (OR) for noncompliance (that is, no helmet worn) attributable to an age-restricted law (in contrast to noncompliance under a universally applicable law) were determined for each studied law change. In the case of SD77, this was a point estimate; for the remaining studies a 95% confidence interval was obtained. In all these remaining studies, at the 5% significance level the OR was found to be greater than one and the APaE exceeded zero.

3.2.2. Heterogeneity analysis

Random-effects models resulted in an I^2 of 53% for each of the APaE and OR statistics. A fixed effect model, after exclusion of the TX89 data, reduced the I^2 to 0% and 17% respectively, without leading to markedly different means (details available from corresponding author).

Pooling of the single-state data using a random-effects model resulted in a mean APaE of 65.3%, with 95% confidence interval of [51.0, 75.4]. A fixed effect model after exclusion of TX89 gave a mean of 62.1% [54.7, 68.3]; an unweighted pooling of single-state data led to a mean of 72.3% [65.7, 78.9]. In lay terms, this means that about two-thirds of noncompliance occurring under an age-restricted helmet law disappears under a universal law. The OR results for these models were 4.41 [3.13, 6.23] (random-effects), 3.86 [3.07, 4.86] (fixed effect), and 6.71 [5.57, 8.08] (unweighted). As an indication, an OR of 4 would correspond to 25% noncompliance with an age-restricted law (odds of 3 to 1) shrinking to 7.7% noncompliance (odds of 12 to 1) after change to a universal law.

3.2.3. Forest plots

Below, we display the forest plots for both attributable percent among exposed (APaE) Fig. 2 and odds ratio (OR) Fig. 3, comparing youthful noncompliance with universal and partial helmet laws. The mean figures for each state, for the weighted pooling of states (other than South Dakota, whose data were too limited to allow pooling), and for each of the national studies, are presented in the third column. The graphical display shows that mean as a small square located within the 95% confidence interval as a horizontal bar. (The size of the square relates inversely to the length of the confidence interval and represents the weighting in the pooled results.) For the pooled results, the confidence interval corresponds to the horizontal endpoints of the kite-like figure, and the mean corresponds to the vertically broadest part of the figure. Note that APaE can theoretically achieve any value between 0 and 100, with higher values representing a higher proportion of noncompliance with an age-specific law that disappears under a universal law regime. The OR can theoretically have any positive value, with values in excess of 1 indicating higher odds of youth noncompliance with an age-restricted helmet law relative to the odds of youth noncompliance with a universal law. The compression induced by the 100 upper bound for APaE makes an outlier like TX89 less obvious in the first forest plot than in the second.

4. Discussion

4.1. Outcomes

4.1.1. Heterogeneity and model tests

Heterogeneity among states may result from cultural differences arising from, for example, climate and geography, history or demography. Differences in outcomes may also reflect the effectiveness and zeal of accompanying campaigns, media attention, political contention over the law-change, and comments by high-profile individuals. Differences in data collection methods may also contribute to heterogeneity among studies. Nevertheless, even for the random-effects model, acceptable levels of I^2 were obtained.

4.1.2. Interpretation of results

As noted by the Community Preventive Services Task Force (2016), commenting on a subset of the studies reported here, "Although each study design comes with unique risks of bias, effect estimates across multiple study types, population groups, and outcome measures were remarkably consistent for this body of evidence. No plausible source of bias could account for this

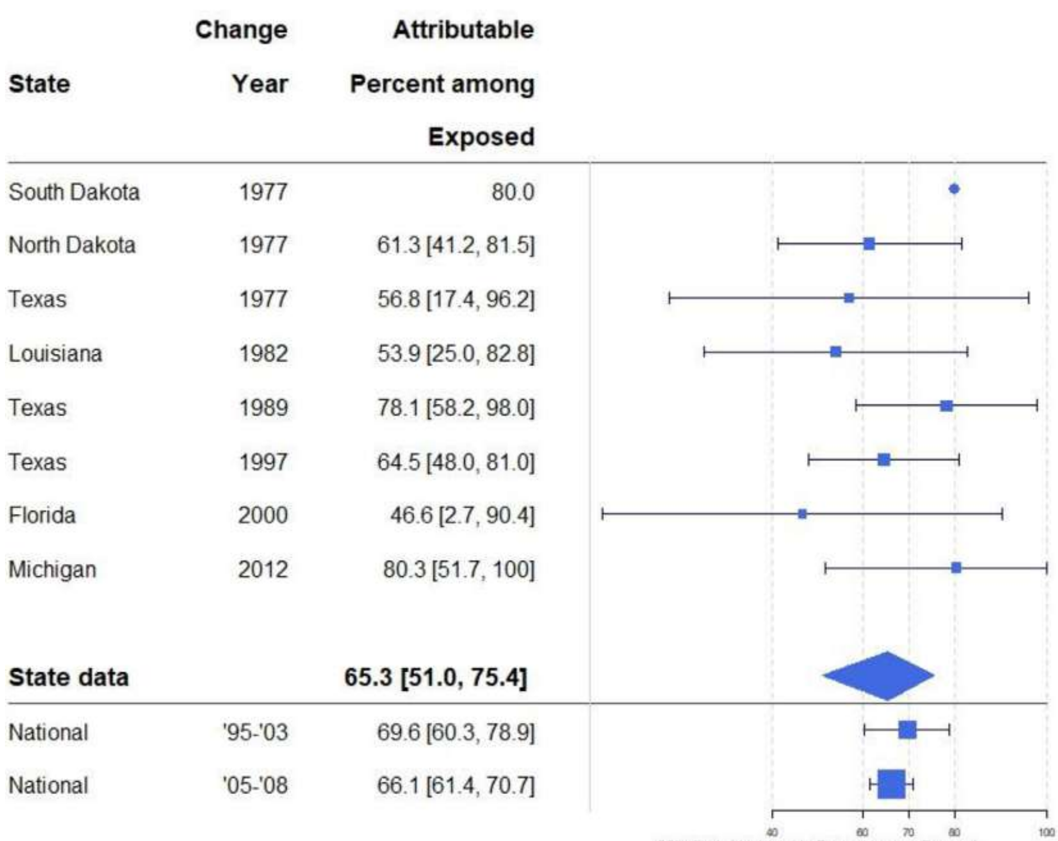


Fig. 2. Young operator noncompliance with age-specific versus universal helmet laws Forest plot: Attributable Percentage among Exposed.

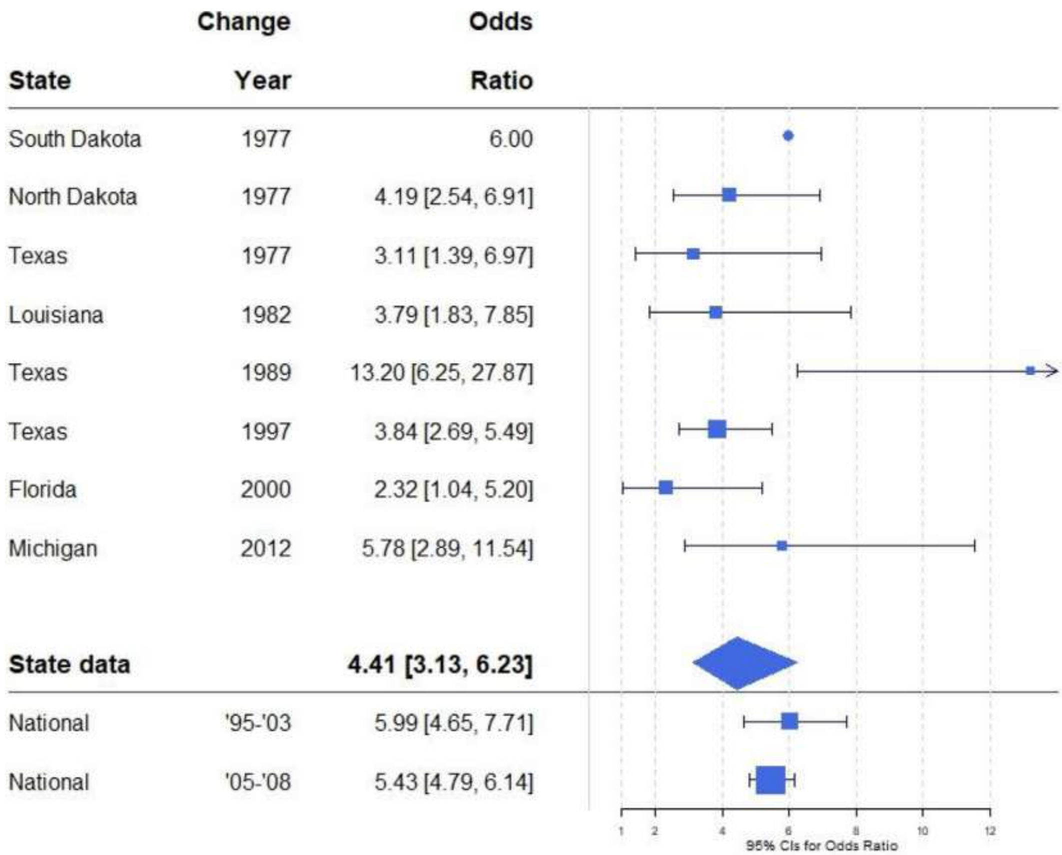


Fig. 3. Young operator noncompliance with age-specific versus universal helmet laws Forest plot: Odds Ratio.

consistency" (ibid, p.11). The highly overlapping confidence intervals obtained from pooled-intrastate and national data for APaE and OR in Table 1 may be taken as a broad indication of the size of this effect. The finding of Brooks et al. (2010) is that "Partial [age-specific] helmet laws neither significantly reduce fatality rates nor increase helmet compliance rates among young riders. A partial helmet law is roughly equivalent to none at all" (ibid, Abstract).

4.2. Limitations

The cited studies rely on three different types of measurement of helmet usage: all riders in crashes, crash injury victims, and fatalities. Since each metric leads to different numbers, there are limitations to comparing figures prior and subsequent to legislative changes by means of different measurement types. Even where there is consistency of metric, each has its own limitations.

Crash or crash injury data involving unhelmeted motorcyclists in breach of the law may be under-reported owing to legal liability and insurance claim issues. Moreover, because of differences in state laws, the extent of such under-reporting may vary among states and thus be very difficult to make accurate allowance for. None has been made in the present analysis. A limitation of fatality data is that the sample is smaller than for injuries. Perhaps differential rates of fatalities over time relate to differences in quality of health care, or changing attitudes and behavior. However, for a comparison of rates in the same state separated by only a few years, these effects are likely to be small. Other differences may relate to demographics, and publicity from the political debate accompanying the law change.

The use of fake/novelty helmets is a complicating issue. Rice et al. (2017) found that riders in Californian collisions wearing novelty helmets were almost twice as likely to die as riders with full-face helmets. The decision to wear such helmets may be related to the type of motorcycle driven and the gender of the driver and passenger (Turner & Hagelin, 2000). However, the relationship between the usage of such helmets and the prevailing law is not clear-cut (Preusser et al., 2000; Turner & Hagelin, 2004).

4.3. Implications

4.3.1. U.S. evidence

There is evidence that drivers are more likely to wear helmets for rural trips (Dorris & Purswell, 1978; Krane & Winterfield, 1980; U.S. GAO, 1991; Gkritza, 2009). There, overall vehicle mileage death rates are much higher (O'Neill & Kyrychenko, 2006), and fear of apprehension for flouting helmet laws is presumably less, than in cities. This would appear to undermine the conspicuity and enforcement prioritization hypotheses presented in Section 1.6 above. Moreover, youth whose main motivation for wearing a helmet is fear of detection would need to drive so as to minimize the risk of being apprehended for an unrelated offense that would reveal their infringement of the underage helmet law. "Enforcement of the [underage] law almost always is secondary, with officers checking the age and/or insurance coverage only after stopping the motorcycle for another traffic violation" (Kyrychenko & McCartt, 2006, p.56). This more cautious behavior would in turn make it less likely that the rider would appear in the fatality or serious incident statistics that comprise these studies. The effect would be to reduce noncompliance with age-restricted laws detected by these crash-based studies. Yet the size of the differences found here casts doubt that that is what has occurred. Kraus et al. (1995) found that the percentage of youth wearing helmets increased in advance of a universal law taking effect – a period during which there would be no increased threat of apprehension. Likewise, Lund et al. (1991, p.578) observed: "Most of the increases [in helmet use] occurred immediately in Texas despite the announcement by some State officials that they would not enforce the law during its first 90 days."

Numerous studies, both within and outside the United States, provide support for the mixed messaging hypothesis. In the United States, Ranney et al. (2010, p.2061) reported: "The variables with the strongest correlation with not always using a helmet were: not believing that wearing a helmet was protective (tetrachoric rho = 0.80); believing that helmets impair sight/hearing (tetrachoric rho = 0.80);" earlier, Allegrante et al. (1980) found that the primary distinction between helmet-wearers and non-wearers was their belief in safety versus comfort-convenience consequences of helmet use.

4.3.2. International evidence

Despite the limitations of international comparisons, we note various international studies that consider whether youths' adherence to helmet laws may be influenced by the messaging they convey about safety as opposed to enforcement threats. In rural Thailand (where enforcement of motorcycle helmet use laws is described as "poorly enforced"), Swaddiwudhipong et al. (1998) found that an education campaign on the protection afforded by helmets produced in intervention villages an APaE for helmet usage noncompliance of 32% and OR of 3.3 when compared with control villages; these statistics are consistent with the U.S. outcomes recorded above. An Iranian study (Haqverdi et al., 2015) reported that "Perception of enforcement on helmet use does not reliably affect helmet use." In rural Malaysia, Sabahiah and Sukor (2014) found noteworthy differences between respondent age groups, with the strongest indicators of helmet usage in the youngest (16–25 years) age group being attitude, perception of others not using a helmet, and perceived danger; while enforcement played a key role only for older age groups. Likewise, in Cambodia, for Bachani et al. (2012) 86% of respondents indicated the lifesaving potential of a helmet as a reason for wearing one, as opposed to 21.7% nominating police fines. An Italian survey (Bianco et al., 2005) of public secondary students aged 14–19, reported that 78.5% opposed helmet laws applicable only for those under 18 years, whereas, despite 70% considering helmets uncomfortable, only 11.7% opposed universal helmet laws.

4.3.3. Other studies

Focus group studies also support the mixed messaging (Germeni et al., 2009) and rite-of-passage (Eliasson et al., 2010) hypotheses. Quantitative studies in the somewhat related area of bicycle helmet laws are also revealing, despite the limitation of change of domain. A systematic review (Karkhaneh et al., 2006) of a dozen bicycle helmet studies concludes that "perhaps the main effect of the legislation is to educate the community/parents that bike helmets are protective, subsequently changing social norms about helmet use and increasing prevalence . . . even in the absence of rigorous enforcement" (ibid, p.81). For example, an Australian survey (Finch, 1996) of 1,240 13- to 17-year-olds found that the majority would wear (mandatory) bicycle helmets to be safe, whereas fewer than 15% stated fear of apprehension as an argument for wearing one.

Recent growth in behavioral economics (more specifically, choice architecture or "nudges") might encourage the hope that suitable campaigns (or adjustments to, for example, the terms and conditions of insurance policies) could lead to greater adherence to helmet laws in states where universal laws are deemed politically infeasible. Nevertheless, this review indicates that it is the law itself that provides a substantial nudge.

5. Conclusion

The studies in this review indicate that around two-thirds of adolescent noncompliance with motorcycle helmet usage laws that are age-restricted tends to disappear in the context of laws that are universally applicable. Studies of youths' reactions suggest that in large part this is because universal laws send a more convincing message that helmets afford protection against injury. It appears that youth are considerably more likely to respect universal laws that send a message of protective benefit than age-specific laws that may be interpreted as signaling authorities' desire for youth control.

6. Practical applications

The meta-analysis provides a fresh perspective on the relative efficacy of universal helmet laws, for young riders in particular. Further studies also offer insight into adolescent psychology, suggesting that alternatives to age-specific laws be considered more widely in the domain of public safety and health.

Declaration of interest statement

The authors of this manuscript declare no conflicts of interest.

Acknowledgements

The authors wish to thank Dr. Yue Ke (Argonne National Laboratory) for assistance with meta-analysis verification, and Dr. Max Gordon (Karolinska Institute) for help with forest plot coding.

References

Abbas, A. K., Hefny, A. F., & Abu-Zidan, F. M. (2012). Does wearing helmets reduce motorcycle-related death? A global evaluation. *Accident Analysis & Prevention, 49*, 249–252.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T.

Allegrante, J. P., Mortimer, R. G., & O'Rourke, T. (1980). Social-psychological factors in motorcycle safety helmet use: Implications for public policy. *Journal of Safety Research*, 12(3), 115–126. Retrieved from =citjournalarticle_25269_22.

Bachani, A. M., Tran, N. T., Sann, S., Ballesteros, M. F., Gnim, C., Ou, A., ... Hyder, A. A. (2012). Helmet use among motorcyclists in Cambodia: A survey of use, knowledge, attitudes, and practices. *Traffic Injury Prevention, 13*(sup1), 31–36.

Berrick, A. J. (2013). The tobacco-free generation proposal. *Tobacco Control, 22*(suppl 1), i22–i26.

Berrick, J. (2018). Adolescent noncompliance with age-specific versus universal US motorcycle helmet laws: systematic review and meta-analysis (PROSPERO International Prospective Register of Systematic Reviews No. 105045). Retrieved from University of York, UK website: https://www.crd.york.ac.uk/PROSPERO/.

Bianco, A., Trani, F., Santoro, G., & Angelillo, I. F. (2005). Adolescents' attitudes and behaviour towards motorcycle helmet use in Italy. *European Journal of Pediatrics, 164*(4), 207–211.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester: Wiley.

Brooks, E., Naud, S., & Shapiro, S. (2010). Are youth-only motorcycle helmet laws better than none at all? *The American Journal of Forensic Medicine and Pathology, 31*(2), 125–129. https://doi.org/10.1097/PAF.0b013e3181c6beab.

Buckley, L., Bingham, C. R., Flannagan, C. A., Carter, P. M., Almani, F., & Cicchino, J. B. (2016). Observation of motorcycle helmet use rates in Michigan after partial repeal of the universal motorcycle helmet law. *Accident Analysis & Prevention, 95*, 178–186. https://doi.org/10.1016/j.aap.2016.07.009.

Carter, P. M., Buckley, L., Flannagan, C. A. C., Cicchino, J. B., Hemmila, M., Bowman, P. J., ... Bingham, C. R. (2017). The impact of Michigan's partial repeal of the universal motorcycle helmet law on helmet use, fatalities, and head injuries. *American Journal of Public Health, 107*(1), 166–172.

Coben, J. H., Steiner, C. A., & Miller, T. R. (2007). Characteristics of motorcycle-related hospitalizations: Comparing states with different helmet laws. *Accident Analysis & Prevention, 39*(1), 190–196.

Cole, P., & MacMahon, B. (1971). Attributable risk percent in case-control studies. *British Journal of Preventive & Social Medicine, 25*(4), 242.
Community Preventive Services Task Force (2016). Motor vehicle-related injury prevention: Use of motorcycle helmets, universal helmet laws Retrieved January 8, 2018, from https://www.thecommunityguide.org/sites/default/files/assets/MVOI-Motorcycle-Helmets-Laws-Mandating-Use_1.pdf.
Dee, T. S. (2009). Motorcycle helmets and traffic safety. *Journal of Health Economics, 28*(2), 398–412.
Dillard, J. P., & Shen, L. (2005). On the nature of reactance and its role in persuasive health communication. *Communication Monographs, 72*(2), 144–168.
Dorris, A. L., & Purswell, J. L. (1978). Impact of Motorcycle Helmet Usage in Oklahoma. Volume I (No. DOT-HS-803-681 Final Rpt.).
Eliasson, M., Johansson, K., & Laflamme, L. (2010). Age and the social meaning of wearing a bicycle helmet among Swedish teenagers. *Injury Prevention, 16*(Suppl 1), A231–A232.
Evans, L., & Frick, M. C. (1988). Helmet effectiveness in preventing motorcycle driver and passenger fatalities. *Accident Analysis & Prevention, 20*(6), 447–458.
Falco, A., Piccirelli, A., Girardi, D., Dal Corso, L., & De Carlo, N. A. (2013). Risky riding behavior on two wheels: The role of cognitive, social, and personality variables among young adolescents. *Journal of Safety Research, 46*, 47–57.
Finch, C. F. (1996). Teenagers' attitudes towards bicycle helmets three years after the introduction of mandatory wearing. *Injury Prevention, 2*(2), 126–130.
Fleming, N. S., & Becker, E. R. (1992). The impact of the Texas 1989 Motorcycle Helmet Law on total and head-related fatalities, severe injuries, and overall injuries. *Medical Care*, 30(9), 832–845. Retrieved from https://www.jstor.org/stable/3765724.
Gefeller, O. (2001). Editorial - The concept of attributable risk in epidemiology: yesterday, today and tomorrow. *Statistical Methods in Medical Research, 10*, 157–158.
Germeni, E., Lionis, C., Davou, B., & Th Petridou, E. (2009). Understanding reasons for non-compliance in motorcycle helmet use among adolescents in Greece. *Injury Prevention, 15*(1), 19–23.
Gkritza, K. (2009). Modeling motorcycle helmet use in Iowa: Evidence from six roadside observational surveys. *Accident Analysis & Prevention, 41*(3), 479–484.
Goodnow, R. K. (1990). Injury severity, medical costs and associated factors for helmeted and unhelmeted motorcyclist crash cases transported to hospitals in Amarillo, Austin, Corpus Christi, and San Antonio, TX. International Motorcycle Safety Conference. Vol 1. Orlando, Florida. Proceedings of the International Motorcycle Safety Conference, 1.
Groeger, J. A. (2006). Youthfulness, inexperience, and sleep loss: The problems young drivers face and those they pose for us. *Injury Prevention, 12*(suppl_1), i19. https://doi.org/10.1136/ip.2006.012070.
Haqverdi, M. Q., Seyedabrishami, S., & Groeger, J. A. (2015). Identifying psychological and socio-economic factors affecting motorcycle helmet use. *Accident Analysis & Prevention, 85*, 102–110.
Heilman, D. R., Weisbuch, J. B., Blair, R. W., & Graf, L. L. (1982). Motorcycle-Related trauma and helmet usage in North Dakota. *Annals of Emergency Medicine, 11*(12), 659–664.
Ho, E. L., & Haydel, M. J. (2004). Louisiana motorcycle fatalities linked to statewide helmet law repeal. *The Journal of the Louisiana State Medical Society, 156*(3), 151–156.
Homer, J., & French, M. (2009). Motorcycle helmet laws in the United States from 1990 to 2005: Politics and public health. *American Journal of Public Health, 99*(3), 415–423.
Houston, D. J. (2007). Are helmet laws protecting young motorcyclists? *Journal of Safety Research, 38*(3), 329–336. https://doi.org/10.1016/j.jsr.2007.05.002.
IIHS (2019). Fatality Facts 2018: Motorcycles and ATVs. Insurance Institute for Highway Safety. Retrieved from: https://www.iihs.org/topics/fatality-statistics/detail/motorcycles-and-atvs.
Jones, M. M., & Bayer, R. (2007). Paternalism & its discontents: Motorcycle helmet laws, libertarian values, and public health. *American Journal of Public Health, 97*(2), 208–217. https://doi.org/10.2105/AJPH.2005.083204.
Karkhaneh, M., Kalenga, J.-C., Hagel, B. E., & Rowe, B. H. (2006). Effectiveness of bicycle helmet legislation to increase helmet use: A systematic review. *Injury Prevention, 12*(2), 76–82.
Kim, C.-Y., Wiznia, D. H., Averbukh, L., Dai, F., & Leslie, M. P. (2015). The economic impact of helmet use on motorcycle accidents: A systematic review and meta-analysis of the literature from the past 20 years. *Traffic Injury Prevention, 16*(7), 732–738. https://doi.org/10.1080/15389588.2015.1005207.
Krane, S. W., & Winterfield, L. A. (1980). *Impact of motorcycle helmet usage in Colorado: A three year study*. Springfield, VA: U.S. National Highway Safety Administration.
Kraus, J. F., Peek, C., & Williams, A. (1995). Compliance with the 1992 California motorcycle helmet use law. *Am J Public Health, 85*(1), 96–99.
Kyrychenko, S. Y., & McCartt, A. T. (2006). Florida's weakened motorcycle helmet law: Effects on death rates in motorcycle crashes. *Traffic Injury Prevention, 7*(1), 55–60. https://doi.org/10.1080/15389580500377833.
Lee, J., Abdel-Aty, M., Wang, J.-H., & Lee, C. (2017). Long-term effect of universal helmet law changes on motorcyclist fatal crashes: Comparison group and empirical Bayes approaches. *Transportation Research Record, 2637*(1), 27–37.
Lee, J. M. (2018). Mandatory helmet legislation as a policy tool for reducing motorcycle fatalities: Pinpointing the efficacy of universal helmet laws. *Accident Analysis & Prevention, 111*, 173–183.
Lee, J. M., & Outlaw, N. L. (2018). The antecedents of mandatory motorcycle helmet legislation enhancing behavior. *Transportation Research Part F: Traffic Psychology and Behaviour, 59*, 330–347.
Lin, M.-R., & Kraus, J. F. (2008). Methodological issues in motorcycle injury epidemiology. *Accident Analysis & Prevention, 40*(5), 1653–1660. https://doi.org/10.1016/j.aap.2008.05.005.
Liu, B. C., Ivers, R., Norton, R., Boufous, S., Blows, S., & Lo, S. K. (2008). Helmets for preventing injury in motorcycle riders. In The Cochrane Collaboration (Ed.), Cochrane Database of Systematic Reviews. https://doi.org/10.1002/14651858.CD004333.pub3.
Lund, A,K., Williams, A.F., & Womack, K.N. (1991). Motorcycle helmet use in Texas. *Public Health Reports*, 106(5), 576-578.
Mayrose, J. (2008). The effects of a mandatory motorcycle helmet law on helmet use and injury patterns among motorcyclist fatalities. *Journal of Safety Research, 39*(4), 429–432. https://doi.org/10.1016/j.jsr.2008.07.001.
Mcswain, N. E., Jr., & Belles, A. (1990). Motorcycle helmets—medical costs and the law. *The Journal of Trauma: Injury, Infection, and Critical Care, 30*(10), 1189–1199.
NHTSA. (1998). Further analysis of motorcycle effectiveness using CODES linked data (No. PB98-163629). National Center for Statistics and Analysis, Washington D.C.: US Department of Transportation.
Njå, O., & Nesvåg, S. M. (2007). Traffic behaviour among adolescents using mopeds and light motorcycles. *Journal of Safety Research, 38*(4), 481–492.
OECD/ECMT Transport Research Centre, European Conference of Ministers of Transport, Organisation for Economic Co-operation and Development, & United Nations (Eds.). (2006). Young drivers: the road to safety. Paris: Organisation for Economic Cooperation and Development ; European Conference of Ministers of Transport.
Olsen, C. S., Thomas, A. M., Singleton, M., Gaichas, A. M., Smith, T. J., Smith, G. A., ... Yeager, D. (2016). Motorcycle helmet effectiveness in reducing head, face and brain injuries by state and helmet law. *Injury Epidemiology, 3*(1), 8–18.
O'Neill, B., & Kyrychenko, S. Y. (2006). Use and misuse of motor-vehicle crash death rates in assessing highway-safety performance. *Traffic Injury Prevention, 7*(4), 307–318.
Peek-Asa, C., & Kraus, J. F. (1997). Estimates of injury impairment after acute traumatic injury in motorcycle crashes before and after passage of a mandatory helmet use law. *Annals of Emergency Medicine, 29*(5), 630–636.
Peng, Y., Vaidya, N., Finnie, R., Reynolds, J., Dumitru, C., Njie, G., ... Shults, R. A. (2017). Universal motorcycle helmet laws to reduce injuries: a community guide systematic review. *American Journal of Preventive Medicine, 52*(6), 820–832.
Preusser, D. F., Hedlund, J. H., & Ulmer, R. G. (2000). Evaluation of motorcycle helmet law repeal in Arkansas and Texas (No. Contract No. DTNH22-97-D-05018). U.S. Department of Transportation, National Highway Traffic Safety Administration.
Ranney, M. L., Mello, M. J., Baird, J. B., Chai, P. R., & Clark, M. A. (2010). Correlates of motorcycle helmet use among recent graduates of a motorcycle training course. *Accident Analysis & Prevention, 42*(6), 2057–2062.
Rice, T. M., Troszak, L., Erhardt, T., Trent, R. B., & Zhu, M. (2017). Novelty helmet use and motorcycle rider fatality. *Accident Analysis & Prevention, 103*, 123–128.
Rockhill, B., Newman, B., & Weinberg, C. (1998). Use and misuse of population attributable fractions. *American Journal of Public Health, 88*(1), 15–19.
Sabahiah, N., & Sukor, A. (2014). Neglecting helmet usage in rural area: Behavioral causal factors according to different age groups. In InCIEC 2013 (pp. 413–422). Springer.
Scott-Parker, B., Hyde, M. K., Watson, B., & King, M. J. (2013). Speeding by young novice drivers: What can personal characteristics and psychosocial theory add to our understanding? *Accident Analysis & Prevention, 50*, 242–250.
Shope, J., & Bingham, C. (2008). Teen driving: Motor-vehicle crashes and factors that contribute. *American Journal of Preventive Medicine, 35*(3), S261–S271.
Stolzenberg, L., & D'Alessio, S.J. (2003). "Born to be wild": the effect of the repeal of Florida's mandatory motorcycle helmet-use law on serious injury and fatality rates. *Evaluation Review*, 27(2), 131-150.
Struckman-Johnson, C., & Ellingstad, V. S. (1980). *Impact of motorcycle helmet law repeal in South Dakota 1976–79: Executive summary*. Vermillion, SD: University of South Dakota.
Swaddiwudhipong, W., Boonmak, C., Nguntra, P., & Mahasakpan, P. (1998). Effect of motorcycle rider education on changes in risk behaviours and motorcycle-related injuries in rural Thailand. *Tropical Medicine & International Health, 3*(10), 767–770.
Texas Department of Public Safety. (1991). Motorcycle helmets: Do they help? Austin: TDPS, 1978. In Highway Safety - Motorcycle Helmet Laws Save Lives and Reduce Costs to Society (pp. 88–89). Washington DC: US General Accounting Office.
Turner, P. A., & Hagelin, C. A. (2000). Novelty helmet use by motorcycle riders in Florida. *Transportation Research Record, 1734*(1), 69–76.
Turner, P.A. & Hagelin, C. A. (2004). Florida motorcycle helmet use observational survey and trend analysis. Center for Urban Transportation Research, College of Engineering, University of South Florida.
Ulmer, R. G., & Northrup, V. S. (2005). Evaluation of the repeal of the all-rider motorcycle helmet law in Florida (No. DOT HS 809 849). Retrieved from National Highway Traffic Safety Administration website: http://www.nhtsa.dot.gov/people/injury/pedbimot/motorcycle/FlaMCReport/pages/Index.htm.
US General Accounting Office, Washington, DC. (1991). Highway Safety - Motorcycle Helmet Laws Save Lives and Reduce Costs to Society (No. GAO/RCED-91-170). Washington DC.
Vaca, F. (2006). Commentary: Motorcycle helmet law repeal: When will we learn... or truly care to learn? Mosby, Inc 11830 Westline Industrial Dr, St Louis, MO 63146-3318 USA.

Watson, G. S., Zador, P. L., & Wilks, A. (1980). The repeal of helmet use laws and increased motorcyclist mortality in the United States, 1975-1978.. *American Journal of Public Health, 70*(6), 579–585. https://doi.org/10.2105/AJPH.70.6.579.
Weiss, H., Agimi, Y., & Steiner, C. (2010). Youth motorcycle-related brain injury by state helmet law type: United States, 2005-2007. *Pediatrics, 126*(6), 1149–1155.
Wong, J.-T., Chung, Y.-S., & Huang, S.-H. (2010). Determinants behind young motorcyclists' risky riding behavior. *Accident Analysis & Prevention, 42*(1), 275–281.

Jon Berrick held positions in mathematics at Oxford and Imperial College London prior to joining National University of Singapore in 1981, where in 2012 he became the inaugural Professor of Science (Mathematics) (Emeritus from 2019) at its Yale-NUS College. Along with his books and numerous research papers in mathematics, he has had research collaborations in neurophysiology and anatomy, in addition to tobacco control – a subject in which he has been lecturing at Yale-NUS. He argues for a tobacco-free generation approach to tobacco eradication. For work on the mathematics of braids, Professor Berrick received Singapore's National Science Award in 2007.

Konstantina Gkritza holds a Diploma from the National Technical University of Athens, Greece; an M.S. from the Virginia Polytechnic Institute and State University, and a Ph.D. from Purdue University; all in Civil Engineering. Before joining Purdue in Fall 2013, she was an assistant professor at Iowa State University and director of the Sustainable Transportation Systems Program. Her research experience lies in the application of economic concepts and statistical methods for efficient, cost-effective, safe, and sustainable transportation systems. Her research on highway safety focuses on motorcycle safety, crash data analysis and evaluation, and the modeling of driver behavior and culture.

Appendix

Preusser et al. (2000) have called attention to the way in which fatality statistics tend to underestimate the rate r of compliance with helmet laws. To see this, let f be the helmeted proportion of all underage fatalities. Then N crashes of sufficient severity as to lead to the death of those without helmets would involve rN riders with helmets (some of whom survive) and $(1 - r)N$ without helmets (none of whom survives). If helmets are effective in saving the lives of proportion e of their wearers in such incidents, after N of these severe events there are among the fatalities $(1 - e)rN$ riders observed with helmets (since erN survive) and also $(1 - r)N$ observed without helmets (since none survives). Thus, the proportion of helmeted riders seen among the fatalities is

$$f = (1 - e)rN/((1 - e)rN + (1 - r)N).$$

From this, we deduce that

$$r = f/(1 - e + ef),$$

which exceeds f because $f < 1$ implies that $1 - e + ef < 1$.

Worked example

Suppose that both before and after a law change 200 fatalities are recorded, with 64 helmeted under an age-restricted law and 128 under a universal law. These raw data give the following contingency table.

	Age-restricted	Universal
Helmet	64	128
No helmet	136	72
Total fatalities	200	200

When converted to crash data according to the formulae above (using $e = 0.36$), the following table is obtained.

	Age-restricted	Universal
Helmet	100	200
No helmet	136	72
Total serious crashes	236	272
	$R_c = 136/236$	$R_p = 72/272$

For this hypothetical example, the APaE is therefore calculated as APaE = $100 \times (R_c - R_p)/R_c = 54.1\%$.

This can be interpreted as 54 percent of the youth noncompliance with an age-specific law being attributable to its age-restriction.

Journal of Safety Research 76 (2021) 218–227

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Analysis of the severity of vehicle-bicycle crashes with data mining techniques

Siying Zhu

School of Civil and Environmental Engineering, Nanyang Technological University, Singapore

ARTICLE INFO

Article history:
Received 22 March 2020
Received in revised form 19 July 2020
Accepted 23 November 2020
Available online 25 December 2020

Keywords:
Safety
Severity
Vehicle-bicycle crashes
Integrated data mining framework
Gradient boosting algorithm

ABSTRACT

Introduction: Although cycling is increasingly being promoted for transportation, the safety concern of bicyclists is one of the major impediments to their adoption. A thorough investigation on the contributing factors to fatalities and injuries involving bicyclist. *Method:* This paper designs an integrated data mining framework to determine the significant factors that contribute to the severity of vehicle-bicycle crashes based on the crash dataset of Victorian, Australia (2013–2018). The framework integrates imbalanced data resampling, learning-based feature extraction with gradient boosting algorithm and marginal effect analysis. The top 10 significant predictors of the severity of vehicle-bicycle crashes are extracted, which gives an area under ROC curve (AUC) value of 0.8236 and computing time as 37.8 s. *Results:* The findings provide insights for understanding and developing countermeasures or policy initiatives to reduce severe vehicle-bicycle crashes.

1. Introduction

Cycling is becoming increasingly popular in recent years as this mode of transportation is healthy, low-cost and environmentally friendly. Many transportation decision makers aim to make cycling a lifestyle in order to support the car-lite vision and enhance the overall livability of the city or country. However, the safety concern on cycling is one of the major impediments to its adoption, as bicyclists are more vulnerable in comparison with auto-mobile occupants. To achieve the long-term goal of alleviating congestion and pollution by engaging more bicyclists, it is necessary to solve commuters' fear of being involved in a crash (Kaplan & Giacomo Prato, 2015).

Enhancing the safety level of bicyclists is a different challenge compared with motorized traffic which has been well studied in the literature. The crashes involving bicyclists are rare and often severe; bicyclists exposure is different from vehicle exposure which is difficult to quantify; and the crash trends of bicyclists are quite distinctive which depend on land use, existing bicycle infrastructure, socio-economic factors, etc. Raihan, Alluri, Wu, and Gan (2019). A thorough investigation on different characteristics contributing to fatalities and injuries involving bicyclists is necessary.

For road safety analysis, the application of non-parametric and data mining technique becomes increasingly popular in recent years, which refers to the analytic process designed to explore big data, searching for structures, commonalities, and hidden patterns or rules (Prati, Pietrantoni, & Fraboni, 2017; Han, Pei, & Kamber, 2011). The data mining techniques can handle large and complicated datasets with relatively short data preparation time and provides satisfiable accuracy (Ding, Chen, & Jiao, 2018).

This paper designs an integrated data mining framework to determine the significant factors which contribute to the severity of vehicle-bicycle crashes based on the crash dataset of Victorian, Australia (2013–2018). The framework integrates imbalanced data resampling, learning-based feature extraction with gradient boosting algorithm and marginal effect analysis to determine the significant contributing factors to vehicle-bicycle crashes.

Specifically, in terms of traffic safety, the main contributions of this paper are elaborated as follows: This paper is dedicated to vehicle-bicycle crash severity modeling to address the safety concerns on bicyclists who are vulnerable road users. In vehicle-bicycle crash severity analysis, the class imbalance issue of the crash dataset exists as the proportion of fatal or severe crashes is relatively small (Prati et al., 2017). The problem can be handled by the imbalanced data resampling process in the integrated data mining framework. The complexity of crash dataset can also be addressed with the learning-based feature extraction process in an iterative manner, in order to determine the most significant contributing factors to the severity of vehicle-bicycle crashes and cater for the trade-off between computation time and model performance. Moreover, the vehicle-bicycle crash dataset in this paper contains a large number of discrete variables. The large number of categories can be handled by gradient boosting algorithm, relying

E-mail address: siying001@e.ntu.edu.sg

https://doi.org/10.1016/j.jsr.2020.11.011
0022-4375/

on no strict statistical assumption (Saha, Alluri, & Gan, 2015; Ding, Cao, & Næss, 2018; Zheng, Lu, & Lantz, 2018). The impact of the most significant contributing factors on the severity of vehicle-bicycle crashes are explained with the marginal effect analysis, and the result from the integrated data mining framework can provide some implications for policies and counter-measures for fatal and serious vehicle-bicycle crashes.

2. Related work

Research on crash severity modeling has been conducted to identify the most significant predictors of the severity of vehicle-bicycle crashes. Klop and Khattak (1999) identified the contributing factors to bicycle crash severity on two-lane, undivided roadways with ordered probit model. Kim, Kim, Ulfarsson, and Porrello (2007) analyzed the determinants of bicyclists injury severity in vehicle-bicycle crashes with multinomial logit model. Yan, Ma, Huang, Abdel-Aty, and Wu (2011), Bahrololoom, Moridpour, and Tay (2016) analyzed the interrelationship of irregular maneuver, crash patterns, etc., and cyclist injury severity with binary logit model. Kaplan, Vavatsoulas, and Prato (2014), Bahrololoom, Moridpour, Tay, and Sobhani (2017) analyzed the determinants of cyclist injury severity level with generalized ordered probit model and generalized ordered logit model. Helak et al. (2017) utilized univariate and multiple regression analyses to study the influence of bike lanes, alcohol, lighting, speed, and helmet on the injury severity of bicyclists. Robartes and Chen (2017) identified the factors that affect cyclist injury severity in the case of single bicycle-single vehicle crashes, in consideration of the cyclist, auto-mobile driver, vehicle, environmental and roadway characteristics with an ordered probit model. Behnood and Mannering (2017) analyzed the factors that significantly affect bicycle injury severities in vehicle-bicycle crashes, utilizing a random parameters multinomial logit model with heterogeneity in means and variances. Bahrololoom, Young, and Logan (2018) investigated the effect of factors related to the three pillars of the Safe System approach including 'safe roads and roadsides', 'safe speeds' and 'safe road users' on bicycle crash severity with random parameter binary logit model. Yasmin and Eluru (2018) analyzed the total crash count and crash proportion by various crash severity levels based on a joint negative binomial-ordered logit fractional split econometric model framework. Sivasankaran and Balasubramanian (2020) applied latent class clustering algorithm to analyse the crash severity of vehicle-bicycle crashes for each cluster. Liu, Khattak, Li, Nie, and Ling (2020) investigated bicyclists injury severity with geographically weighted ordinal logistic regression model to address the spatial heterogeneity.

In particular, research has been conducted for crash severity modeling of vehicle-bicycle crashes on specific types of infrastructure, such as bike lanes and intersections. Klassen, El-Basyouny, and Islam (2014) investigated the severity level of vehicle-bicycle intersection-related and mid-block-related crashes with spatial mixed logit model. Wall et al. (2016) evaluated the influence of sharrow, painted bicycle lane and physically protected path on bicyclist injury severity with negative binomial model. Moore, Schneider, Savolainen, and Farzaneh (2011) examined bicyclists injury severity in vehicle-bicycle crashes at intersection and non-intersection respectively with mixed logit model. Wang, Lu, and Lu (2015) analyzed the contributing factors bicyclists' injury severity level in vehicle-bicycle crashes at unsignalized intersections with partial proportional odds model. Stipancic, Zangenehpour, Miranda-Moreno, Saunier, and Granié (2016) investigated the contributing factors to the severity of vehicle-bicycle conflicts at urban intersections with ordered logit model based on video data and post-encroachment time, taking the gender differences into account. Asgarzadeh, Verma, Mekary, Courtney, and Christiani (2017) analyzed the effects of intersection and street design on vehicle-bicycle crash severity with multivariate log-binomial regression model, and the results indicated that non-orthogonal intersections and non-intersection segments are associated with higher crash severity. Bahrololoom, Young, and Logan (2018), Bahrololoom, Young, and Logan (2018) investigated the impact of kinetic energy on crash severity of bicyclists at intersections. Rash-ha Wahi, Haworth, Debnath, and King (2018) studied the effects of various traffic control types at intersection with the application of separate mixed logit models for bicyclist injury severity.

Ordinal regression models has been commonly applied to formulate the crash severity model since the injury outcomes are ordinal from no injury to fatal. More recently, to address the limitation of the assumption that all parameters estimated in the model are constant across observations and the heterogeneity of the crash outcomes, some multi-nomial logit models and mixed logit models have been applied for crash severity modeling (Li, Ma, Zhu, Zeng, & Wang, 2018). Regression models relies on strict statistical assumptions, for example, linearity in modeling the relation, which can hardly be satisfied in most crash circumstances. Moreover, the performance of the regression model is poor when handling mass complicated crash data with many discrete variables or variables with a large number of categories satisfactorily (Prati et al., 2017; Li et al., 2018; Ding et al., 2018). To overcome the shortcomings of statistical models, data mining techniques which examines the pre-existing large database have been applied for crash severity modeling. Prati et al. (2017) applied the CHi-squared Automatic Interaction Detection (CHAID) decision tree and Bayesian network analysis to predict the severity of bicycle crashes corresponding to the factors related to crash characteristics. The Bayesian network analysis was further applied to identify the most significant predictors. However, when the complexity of the crash dataset gets larger, feature extraction process is necessary to be applied to address trade-off between computing time and model performance.

Based on the literature review, the research gaps in terms of traffic safety are summarized as follows: (1) In the field of traffic safety, limited research was dedicated to crash severity modeling for vehicle-bicycle crashes in comparison with vehicle-vehicle crashes. (2) The crash severity levels in the vehicle-bicycle crash dataset is highly imbalanced, affecting the performance of crash severity classification model. (3) The data mining techniques which can overcome some shortcomings of statistical models such as the reliance on strict statistical assumptions have rarely been used for the analysis of crash severity of vehicle-bicycle crashes. (4) The learning-based feature extraction process has not been applied in the literature to address the complexity of the crash dataset.

This paper aims to address the research gap by determining the significant factors which contribute to the severity of vehicle-bicycle crashes with a data mining framework, integrating imbalanced data resampling, learning-based feature extraction and marginal effect analysis. The framework introduced in this paper uses gradient boosting as the key algorithm for feature extraction, which can handle different types of predictor attributes, require little data preprocessing effort, and can fit complex nonlinear relationship (Elith, Leathwick, & Hastie, 2008; Zhang & Haghani, 2015; Zheng et al., 2018).

3. Methodology

3.1. Integrated data mining framework

An integrated data mining framework is designed to extract the key crash-related features and predict vehicle-bicycle crash severity level, which integrates imbalanced data resampling, learning-based feature extraction, and marginal effect analysis. The framework is illustrated in Algorithm 1.

Algorithm 1: Integrated data mining framework

Input: Vehicle-bicycle crash severity dataset D
Output: Vehicle-bicycle crashes severity level prediction
Step 1: Resample imbalanced dataset
(a) Synthetic Minority Over-sampling Technique (SMOTE)
(b) Resampled crash dataset D'
Step 2: Learning-based feature selection
(a) Train gradient boosting model based on D'
(b) Determine relative feature importance
(c) Recursive feature elimination
(d) Dataset with key vehicle-bicycle crash features
Step 3: Marginal effect analysis

As the vehicle-bicycle crash dataset is with highly imbalanced classes, data resampling process is applied. For learning-based feature selection, gradient boosting algorithm which is an ensemble learning technique is applied as the key algorithm. The key features for bicycle-vehicle severity prediction is extracted recursively, in consideration of the trade-off between AUC (area under the receiver operating characteristic curve) value and computing time. The receiver operating characteristics (ROC) is a probability curve which demonstrates a comparison of two operating characteristics, namely, specificity and sensitivity, as the threshold changes (Beshah & Hill, 2010). Then, the AUC value serves as a measure of separability, quantifying the overall capability of the model in distinguishing between classes (Narkhede, 2018; Bradley, 1997). The AUC value of 0.5 represents an entirely random test, while the AUC value of 1 represents a perfect classification test. The features extracted are further included in marginal effect analysis to investigate their impact on the vehicle-bicycle crash occurrences.

3.2. Resample imbalanced data

The vehicle-bicycle crash dataset is imbalanced as the classes of crash severity levels are not approximately equally represented. Although predictive accuracy is commonly applied to evaluate the performance of machine learning techniques, it is not suitable when the dataset is imbalanced and the cost of different errors varies greatly. In such situations, the Receiver Operating Characteristic (ROC) curve typically serves as the performance measurement to optimize, which is a probability curve calculated by the true positive rate on the y-axis against the false positive rate on the x-axis. Every point on the ROC curve corresponds to a pair of sensitivity and specificity values based on specific decision threshold. The area under the ROC curve (AUC) serves as a measure of separability between the two classes of crash severity in this paper, such that the value closer to 1 indicates a better separability of model. This paper applies Synthetic Minority Over-sampling Technique (SMOTE) to resample the original crash dataset, which synthesises new minority instances between existing minority instances instead of over-sampling to replacement (Chawla, Bowyer, Hall, & Kegelmeyer, 2002). Moreover, the majority instances are also under-sampled, whose output is a more balanced dataset.

3.3. Gradient boosting algorithm & learning-based feature selection

The gradient boosting algorithm is a type of ensemble learning technique, which sequentially fits a simple parameterized function or base learner into current 'pseudo'-residuals by least squares in each iteration in order to construct additive regression models. At each step, the weightage of observations is adjusted as the subsequent predictors learn from the mistakes of previous predictors. Regarding the model values at each training data point evaluated in the current step, the pseudo-residuals are the gradient of the loss functions being minimized (Friedman, 2002).

Let N be the total number instances in the dataset, and M be the total number of trees to be generated. Let x be the feature vector with a set of predictors and $F(x)$ be an approximation function of the target variable (i.e. severity level of vehicle-bicycle crashes). The gradient boosting algorithm estimates the function $F(x)$ as an additive expansion based on the base learner function $b(x;a_m)$ (Ding et al., 2018; De'Ath, 2007; Saha et al., 2015; Chung, 2013; Zhang & Haghani, 2015):

$$F(x)=\sum_{m=1}^{M}F_m(x)=\sum_{m=1}^{M}\lambda_m b(x;a_m) \quad (1)$$

$$b(x;a_m)=\sum_{d=1}^{D}\gamma_{dm}I(x\in R_{dm}) \quad (2)$$

where each decision tree m divides the input space into D disjoint regions $R_{1m},\dots,R_{Dm}$ and predicts a constant value γ_{dm} for each region R_{dm};

$$I(x\in R_{dm})=\begin{cases}1, & if\ x\in R_{dm}\\ 0, & otherwise\end{cases} \quad (3)$$

a_m represents the mean of split locations and the leaf node for each splitting variable in tree m; λ_m represents weights given to the nodes of decision tree and determines how predictions from the individual decision trees are combined. λ_m is calculated by minimizing a specified loss function, which is a squared error function:

$$L(y,F(x))=(y-F(x))^2 \quad (4)$$

The gradient boosting model is built in a stage-wise fashion, which is updated by minimizing the expected value of the loss function. To avoid over-fitting and improve accuracy, the learning rate or shrinkage, is used to scale the contribution of each base tree learner by introducing a factor of $\varepsilon(0<\varepsilon\leqslant 1)$ as below (Ding et al., 2018; Friedman, 2002):

$$F_m(x)=F_{m-1}(x)+\varepsilon\cdot\lambda_m b(x;a_m) \quad (5)$$

A hybrid two-step learning-based feature selection model is applied to select the key features for vehicle-bicycle crash severity modelling, which is summarized in Algorithm 2. The procedure firstly train and tune the model to rank the feature importance, then the process is permuted to determine an optimal subset of features with Recursive Feature Elimination (RFE), in consideration of the trade-off between the area under ROC curve and computing time.

Algorithm 2: Learning-based feature selection

Input: Crash severity dataset with the set of all predictors as S
Output: Model corresponding to subset S^* of predictors
\\Feature importance ranking:
Train the model with all predictors
Calculate model performance
Rank predictors according to feature importance
\\RFE:
for Subset $S_i \in S$ $(i = 1, ..., |S|)$ **do**
  Keep i most important predictors
  Train the model based on the subset of predictors S_i
  Calculate model performance
end
Generate the performance profile over all subsets of predictors S_i
Select the appropriate number of predictors S^*

3.4. Marginal effect analysis

The sensitivity analysis of traditional linear regression models can only evaluate one predictor at one time such that it ignores the correlation among other predictors. In this paper, the non-linear effects of a set of predictors on the severity of vehicle-bicycle crashes can be illustrated with partial dependence plots generated based on the integrated data mining framework. The partial dependence plot of each predictor demonstrates its marginal effect on the target variable in consideration of the average influences of all other predictors (Saha et al., 2015; Ding et al., 2018).

3.5. Data description

The crash dataset for model estimation comes from Victoria police crash reports across the entire state included 6237 crashes involving bicyclists and motorists between 2013 and 2018 in the entire Victoria State in south-eastern Australia (VicRoads, 2019). Victoria is the smallest mainland state and the second densely populated state in Australia. All crashes included in the analysis are finished cases, while reopened cases are excluded. The attributes of dataset which describe the characteristics of the vehicle-bicycle crashes are summarized in Table 1. Vehicle-bicycle crash injury severity is the dependent variable, which is classified into three categories, namely, fatal accident, serious injury accident and other injury accident. As the percentage of fatal accident is extremely low in comparison with the other types of injuries, the injury severity is finally categorized as Fatal and serious injury accident (1), Other injury accident(0). 30 independent variables are selected for analysis which are classified into five categories, including the type of accident, time factor, characteristics of vehicles, characteristics of environment condition, and human factors. The definitions of some variables are further clarified as below: The road condition classification is based on the State-wide Route Numbering Scheme (SRNS) (data.vic, 2019): 'M' represents the roads which provide a consistent high standard of driving conditions, with divided carriageways, four traffic lanes, sealed shoulders and line marking easily visible in all weather conditions; 'A' represents the roads with similar high standard of driving conditions on a single carriageway; 'B' represents the sealed roads, which are wide enough for two traffic lines, with good centre line and edge line marking, shoulders, and a high standard of guidepost delineation; 'C' represents the roads that are generally two lane sealed with shoulders; others are not classified. Different types of collisions are defined based on Victoria State's local definitions for classifying accidents (DCA code) (VicRoads, 2013).

4. Results analysis

4.1. Model optimization

The vehicle-bicycle crash dataset is randomly separated into two sets, namely training set (80%) and test set (20%). To optimize the model, this paper applied a ten-fold cross-validation procedure which repeated three times on the training set to determine the optimal combination of parameters. The training set is randomly partitioned into ten sub-samples, and each of them is used as the test set while the remaining sub-samples serve as the training set.

The parameters which have been tuned from grid search based on AUC values are explained as below: The shrinkage value or learning rate is introduced to reduce the influence of each individual tree structure and leave space for future trees for improving the model (Friedman, 2002), which is set as 0.1. The lower learning rate can lead to longer computation time. The number of trees indicate the number of gradient boosting iterations, and if the value is too high, it may lead to overfitting, and the value is set as 3. The interaction depth indicate the maximum number of splits of each tree, which is set as 150. The minimum number of observations in the terminal nodes of the trees is also tuned, which is set as 10.

4.2. Recursive feature selection and importance ranking

After model training, recursive feature selection process has been carried out in consideration of the trade-off between AUC value and computing time, as shown in Fig. 1a. The relative contribution rankings or relative importance of the twenty most significant explanatory variables in predicting the severity of vehicle-

Table 1
Descriptive statistics.

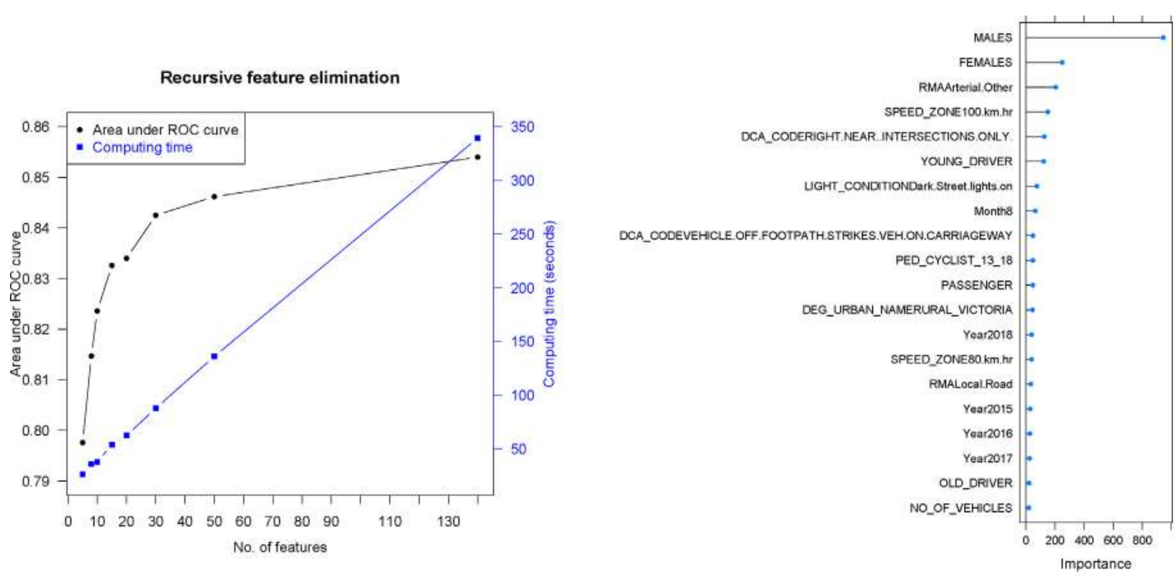
Category	Variable	Count	%
Severity	Fatal accident	33	0.53
	Serious injury accident	1467	23.52
	Other injury accident	4737	75.95
	Accident type		
Collision type	Right through	994	16.01
	Cross traffic	793	12.78
	Vehicle strikes door of parked/stationary vehicle	714	11.5
	Left turn side wipe	458	7.38
	Vehicle off footpath strikes another vehicle while emerging from driveway	413	6.65
	Vehicle strikes another vehicle while emerging from driveway	408	6.57
	Left near	359	5.78
	Right near	308	4.96
	Rear end	279	4.49
	Lane side swipe	243	3.91
	Out of control on carriageway	163	2.63
	Entering parking	108	1.74
	Right far	84	1.35
	Y turn	71	1.14
	Right turn side swipe	65	1.05
	Head on	63	1.01
	...	...	...
Alcohol related	No	6217	99.68
	Yes	20	0.32
	Time factor		
Year	2013	584	9.36
	2014	1298	20.81
	2015	1221	19.58
	2016	1104	17.70
	2017	1054	16.90
	2018	976	15.65
Month	January	442	7.09
	February	547	8.77
	March	683	10.95
	April	499	8.00
	May	517	8.29
	June	397	6.37
	July	477	7.65
	August	481	7.71
	September	437	7.01
	October	626	10.04
	November	563	9.03
	December	568	9.11
Day of week	Monday	883	14.39
	Tuesday	1044	17.02
	Wednesday	1081	17.62
	Thursday	1081	17.62
	Friday	881	14.36
	Saturday	582	9.49
	Sunday	583	9.50
	Vehicle characteristics		
No. of vehicles involved	2	5987	95.99
	3	226	3.62
	4	11	0.18
	5	5	0.08
	6	4	0.06
	7	2	0.03
	8	1	0.02
	14	1	0.02
No. of heavy vehicles involved	0	6143	98.49
	1	94	1.51
No. of passenger vehicle involved	0	224	3.59
	1	5878	94.24
	2	126	2.02
	3	3	0.05
	4	1	0.02
	5	3	0.05
	6	1	0.02
	13	1	0.02
No. of public vehicles	0	6180	99.09
	1	57	0.91
Involve vehicle run off-road	No	14210	95.87
	Yes	206	1.43

	Environment condition characteristics		
Light condition	Dark no street lights	40	0.64
	Dark street lights off	3	0.05
	Dark street lights on	647	10.37
	Dark street lights unknown	83	1.33
	Day	4392	70.42
	Dusk/Dawn	865	13.87
Road Geometry	Cross intersection	1914	25.87
	Multiple intersection	137	2.20
	Not at intersection	2237	35.87
	T intersection	1922	30.82
	Y intersection	16	0.26
	Dead end	2	0.03
Speed zone	110 km/h	3	0.05
	100 km/h	93	1.49
	90 km/h	7	0.11
	80 km/h	249	3.99
	75 km/h	1	0.02
	70 km/h	195	3.13
	60 km/h	2391	38.34
	50 km/h	1579	25.32
	40 km/h	1149	18.42
	30 km/h	17	0.27
	Campus ground or off-road	59	10.95
	Other speed limit	17	0.27
Node type	Intersection	3823	61.31
	Non-intersection	2387	38.28
	Off-road	26	0.42
Urbanized area	Melbourne Urban	4983	79.89
	Small cities	315	5.05
	Melbourne CBD	303	4.86
	Large provincial cities	276	4.43
	Rural Victoria	197	3.16
	Towns	148	2.37
	Small towns	15	0.24
Road classification	Arterial highway	657	10.75
	Arterial other	2378	39.91
	Freeway	48	0.79
	Local road	3028	49.55
Road condition by the Statewide Route Numbering Scheme	A	80	12.82
	B	125	20.03
	C	369	59.13
	M	50	8.01
Crash occur on a divided portion of road	Divided	4324	70.76
	Undivided	1787	29.24
Where crash occur	Metro region	5426	87
	Country region	811	15.86
	Demographics		
Hit-and-run	No	5911	94.77
	Yes	326	5.23
No. of males involved	0	655	10.50
	1	2772	44.44
	2	2580	41.37
	3	193	3.09
	4	22	0.35
	5	14	0.22
	7	1	0.02
No. of females involved	0	2705	43.37
	1	2750	44.09
	2	724	11.61
	3	41	0.66
	4	10	0.16
	5	5	0.08
	6	1	0.02
	7	1	0.02
No. bicyclists involved	1	6146	98.54
	2	79	1.27
	3	6	0.10
	4	4	0.06
	5	1	0.02
	6	1	0.02
No. of pedestrians involved	0	6227	99.84
	1	9	0.14
	4	1	0.02
No. of drivers involved	1	6127	98.24
	2	103	1.65
	3	3	0.05
	4	1	0.02

(*continued on next page*)

Table 1 (*continued*)

Category	Variable	Count	%
	5	2	0.03
	7	1	0.02
No. of vehicle passengers involved	0	5601	89.80
	1	509	8.16
	2	78	1.25
	3	31	0.50
	4	15	0.24
	5	3	0.05
No. of 5–12 year old cyclists involved	0	6043	96.89
	1	1	191
	2	3	0.05
No. of 13–18 year old cyclists involved	0	6236	99.98
	1	1	0.02
No. of 65 years and older pedestrians involved in the crash	0	5929	95.06
	1	308	
No. of 65 years and older drivers involved	0	5929	95.06
	1	308	
No. of 18–25 year old young drivers involved	0	5448	87.35
	1	787	12.62
	2	2	0.03
Unlicensed driver	0 (no)	6155	98.69
	1 (yes)	82	1.31



(a) Trade-off of RFE performance (b) Feature importance ranking

Fig. 1. Learning-based feature selection.

bicycle crashes are summarized with Fig. 1b. The calculated importance ranking scores demonstrate the association between the crash-related predictors and vehicle-bicycle crash severity level. The higher the score, the more significant the predictor. 140 vehicle-bicycle crash features are included in the analysis after sparsifying the crash dataset. As the number of features increases from 5 to 140, the AUC value increases with decreasing gradient from 0.7976 to 0.8540, while the computing time also increases from 26.3 s to 339.4 s.

For further analysis, the top ten predictors have been selected, including the number of males and females involved in the crash, the road type is arterial (other than highway), the speed zone of 100 km/h, the number of young drivers involved in the crash, the light condition (dark with street lights on), the month (August), the collision type (right turn near intersections,bicycle off footpath strikes the vehicle on the carriageway), the number of bicyclists from 13 to 18 years old, with AUC value as 0.8236 and computing time as 37.8 s.

4.3. Marginal effects of key predictors

The partial dependence plots of the significant predictors on the severity of vehicle-bicycle crashes has been plotted for marginal effect analysis, where the relative logit contribution of the predic-

Table 2
Top ten predictors.

No.	Predictor	Effect
1	No. of males involved in the crash	Positive
2	No. of females involved in the crash	Positive
3	Road classification (arterial other)	Positive
4	Speed zone (100 km/h)	Positive
5	Collision type (right turn near intersections)	Positive
6	No. of young drivers involved in the crash	Negative
7	Light condition (dark with street lights on)	Positive
8	Month (August)	Positive
9	Collision type (bicycle off footpath strikes the vehicle on the carriageway)	Positive
10	No. of bicyclists from 13 to 18 years old	Negative

tor on the class probability of 'Fatal & serious injury accident' is plotted as the y-axis. More detailed information on the concept of partial dependence plot can be found in Friedman (2002). The general positive or negative effects of the top ten predictors on the vehicle-bicycle crash severity level reflected by the partial dependence plots are described and summarized in Table 2.

4.4. Discussion

4.4.1. Demographics

As for the impact of demographics, the result shows that the crash is more likely to be a fatal or severe injury accident as the number of males involved in the accident increases. The result is consistent with Kim et al. (2007), Eluru, Bhat, and Hensher (2008), Behnood and Mannering (2017), which suggested that male bicyclists are more likely to be involved in severer crashes. In this paper, we also found that as the number of females involved in the crash increases, the crash is more likely to be severe or fatal. However, the impact of the number of females involved in the crash is less significant than the number of males involved in the crash on the prediction of vehicle-bicycle crash severity.

On the other hand, the results also show that the vehicle-bicycle crash is less likely to be fatal or severe when young drivers or bicyclists are involved in the crash. The result is consistent with the literature (Prati et al., 2017; Yan et al., 2011; Bíl, Bílová, & Müller, 2010), which suggested that bicyclists' injury severity increases with age. The result can be explained by the physical fragility of elder bicyclists, longer perception and reaction time during collision and more inattentive during cycling. It can also be explained that the elder drivers are more susceptible to injury and the higher crash involvement rate due to unsafe driving (Li, Braver, & Chen, 2003). Therefore, more attention should be directed to elder bicyclists' group and driver group to alleviate the vehicle-bicycle crash severity level.

4.4.2. Environment condition characteristics

The results illustrate that vehicle-bicycle crashes on arterials (not highway) are more like to be fatal or serious injury accidents, hence more attention should be paid to developing effective countermeasures on arterials (not highway) to improve the safety level of bicyclists. It is also found that the vehicle speed zone of 100 km/h are more likely to result in fatalities or serious injuries in comparison with other speed zones. The finding is consistent with Robartes and Chen (2017), Kim et al. (2007) which suggested that high speed can significantly affect the severity level of bicyclist-vehicle crashes. It is also found that when the light condition is dark with street lights on, the likelihood of severe and fatal vehicle-bicycle crash increases. This is consistent with the literature, as poor light condition were likely to be associated with more severe consequences of bicycle crashes due to the limited range of visibility (Eluru et al., 2008; Prati et al., 2017).

Although (Robartes & Chen, 2017; Liu et al., 2020) suggested that crashes are less likely to be severe at intersections, the intersection type is not found to be important to crash severity level in this paper. This is understandable, as the intersections still represent major conflict points for bicyclists, despite increasing alertness of bicyclists and drivers at intersections. This may also be explained by the increased numbers of cycle lanes constructed, which can reduce the sensitivity of intersection type to crash severity. In addition, the change in time and space may also cause the difference in the relationship between crash severity and the influential factors (Liu, Hainen, Li, Nie, & Nambisan, 2019).

4.4.3. Time factor

The vehicle-bicycle crash is more likely to be serious or fatal in the August, which is the winter season in Australia. According to Liu, Shen, and Huang (1995), Kaplan and Prato (2013), bicycle crashes is influenced by season and weather conditions, and the result in this paper can be explained by the unpredictability of the weather condition in the specific month.

Unlike (Behnood & Mannering, 2016), year is not identified as a significant contributing factor to crash severity. This may be explained by the fact that (Behnood & Mannering, 2016) utilized the eight-year (2005–2012) crash dataset, covering pre-recession, recession and post-recession period, such that the significant temporal instability was found. On the other hand, this paper only utilized crash dataset for post-recession period such that the long-term effect of time factor is not apparent.

4.4.4. Accident type

Previous research has also addressed the impact of collision types on crash severity level (Kim et al., 2007; Bíl et al., 2010; Yan et al., 2011; Behnood & Mannering, 2017; Prati et al., 2017). In this paper, it is found that when the vehicle turns right near intersections and when bicycle off footpath strikes the vehicle on the carriageway, the vehicle-bicycle crash severity is more likely to increase. The result can be explained by the unexpected event and higher level of kinetic energy during the particular types of collision in comparison with others.

Although head-on vehicle-bicycle collision or facing the traffic was found to be important in Kim et al. (2007), Liu et al. (2020), it is not identified as a significant predictor to crash severity in this paper. This can be explained by the fact that head-on interaction causes higher relative speed as well as more rapid response of drivers and bicyclists since they are able to see their conflict party before collision. Higher relative speed may increase the crash severity while more proactive reaction may reduce it. Moreover, the type of opponent vehicles involved in the vehicle-bicycle crash is not identified as a significant contributing factor to crash severity level in this study, even though it was identified as a significant predictor in Robartes and Chen (2017), Yan et al. (2011). The result can be explained by the difference in the dataset used for analysis and the way of feature extraction.

5. Conclusions

In this paper, an integrated data mining framework which includes imbalanced data resampling, learning-based feature extraction and marginal effect analysis is designed to determine the significant factors contributing to the severity level of vehicle-bicycle crashes based on the crash dataset of Victoria, Australia from year 2013 to 2018.

This paper has been dedicated to crash severity modeling for vehicle-bicycle crashes which has been less commonly addressed in the literature in comparison with vehicle-vehicle crashes. The learning-based feature selection technique based on gradient boosting algorithm has been applied for key feature extraction on the mass complicated crash dataset which contains a large amount of categorical variables, empty entries, etc. The most significant predictors that affect the severity of vehicle-bicycle crashes are extracted which include the number of males and females involved in the crash, the road type is arterial (other than highway), the speed zone of 100 km/h, the number of young drivers involved in the crash, the light condition (dark with street lights on), the month (August), the collision type (right turn near intersections, bicycle off footpath strikes the vehicle on the carriageway), the number of bicyclists from 13 to 18 years old.

Although cycling is being increasingly promoted as transportation mode, the safety concerns is one of the main obstacles to their adoption. It is a challenge to enhance the safety level of bicyclists to solve the major fear of them, as their safety depends on various factors, including land use, socio-economic factors, etc. The findings of this paper highlight the necessity to identify the contributing factors to fatal and serious vehicle-bicycle crashes to enhance the safety level and guide safety improvements. Several recommendations have been made to improve the safety level of bicyclists. More attention should be paid to the road types and speed zone that are determined to be more prone to severe vehicle-bicycle crashes. For the road user groups, collision types, and time period which are identified to be significant contributing factors to the fatal and serious injury crashes, targeted education campaign should be carried out to enhance the road safety level. The method of this paper can also be extended and applied to other datasets with higher dimensionality for feature selection in order to extract the most significant features and reduce the computation time. Further study can also be carried out to study the crash severity and frequency together and identify the contributing factors to various types of crashes.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Asgarzadeh, M., Verma, S., Mekary, R. A., Courtney, T. K., & Christiani, D. C. (2017). The role of intersection and street design on severity of bicycle-motor vehicle crashes. *Injury Prev., 23*, 179–185.

Bahrololoom, S., Moridpour, S., Tay, R., 2016. Factors affecting bicycle fatal and serious injury crashes in victoria, Australia, in: The 38th Australasian Transport Research Forum, Australian Transport Research Forum, pp. 1–12. .

Bahrololoom, S., Moridpour, S., Tay, R., Sobhani, A., 2017. Exploring the Factors Affecting Bicycle Crash Severity in Victoria, Australia. Technical Report.

Bahrololoom, S., Young, W., Logan, D., 2018a. Exploring the effect of kinetic energy on bicycle crash severity at intersections, in: ARRB International Conference, 28th, 2018, Brisbane, Queensland, Australia. .

Bahrololoom, S., Young, W., Logan, D., 2018b. The role of kinetic energy in bicyclist's injury severity at intersections, in: Australasian Road Safety Conference, 2018, Sydney, New South Wales, Australia. .

Bahrololoom, S., Young, W., Logan, D., 2018c. A Safe System Based Investigation of Factors Influencing Bicycle Crash Severity in Victoria, Australia. Technical Report.

Behnood, A., & Mannering, F. (2017). Determinants of bicyclist injury severities in bicycle-vehicle crashes: a random parameters approach with heterogeneity in means and variances. *Anal. Methods Accid. Res., 16*, 35–47.

Behnood, A., & Mannering, F. L. (2016). An empirical assessment of the effects of economic recessions on pedestrian-injury crashes using mixed and latent-class models. *Anal. Methods Accid. Res., 12*, 1–17.

Beshah, T., Hill, S., 2010. Mining road traffic accident data to improve safety: role of road-related factors on accident severity in Ethiopia, in: 2010 AAAI Spring Symposium Series. .

Bíl, M., Bílová, M., & Müller, I. (2010). Critical factors in fatal collisions of adult cyclists with automobiles. *Accid. Anal. Prev., 42*, 1632–1636.

Bradley, A. P. (1997). The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recogn., 30*, 1145–1159.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). Smote: synthetic minority over-sampling technique. *J. Artif. Intell. Res., 16*, 321–357.

Chung, Y. S. (2013). Factor complexity of crash occurrence: an empirical demonstration using boosted regression trees. *Accid. Anal. Prev., 61*, 107–118.

data.vic, 2019. Crashes Last Five Years. http://data.vicroads.vic.gov.au/metadata/Crashes_Last_Five_Years%20-%20Open%20Data.html. Online; accessed 22 Apr 2019. .

De'Ath, G. (2007). Boosted trees for ecological modeling and prediction. *Ecology, 88*, 243–251.

Ding, C., Cao, X. J., & Næss, P. (2018). Applying gradient boosting decision trees to examine non-linear effects of the built environment on driving distance in oslo. *Transp. Res. A Policy Pract., 110*, 107–117.

Ding, C., Chen, P., & Jiao, J. (2018). Non-linear effects of the built environment on automobile-involved pedestrian crash frequency: a machine learning approach. *Accid. Anal. Prev., 112*, 116–126.

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *J. Anim. Ecol., 77*, 802–813.

Eluru, N., Bhat, C. R., & Hensher, D. A. (2008). A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accid. Anal. Prev., 40*, 1033–1054.

Friedman, J. H. (2002). Stochastic gradient boosting. *Comput. Stat. Data Anal., 38*, 367–378.

Han, J., Pei, J., & Kamber, M. (2011). *Data Mining: Concepts and Techniques*. Elsevier.

Helak, K., Jehle, D., McNabb, D., Battisti, A., Sanford, S., & Lark, M. C. (2017). Factors influencing injury severity of bicyclists involved in crashes with motor vehicles: Bike lanes, alcohol, lighting, speed, and helmet use. *Southern Med. J., 110*, 441–444.

Kaplan, S., & Giacomo Prato, C. (2015). A spatial analysis of land use and network effects on frequency and severity of cyclist–motorist crashes in the copenhagen region. *Traffic Injury Prev., 16*, 724–731.

Kaplan, S., & Prato, C. G. (2013). Cyclist–motorist crash patterns in denmark: a latent class clustering approach. *Traffic Injury Prev., 14*, 725–733.

Kaplan, S., Vavatsoulas, K., & Prato, C. G. (2014). Aggravating and mitigating factors associated with cyclist injury severity in denmark. *J. Saf. Res., 50*, 75–82.

Kim, J. K., Kim, S., Ulfarsson, G. F., & Porrello, L. A. (2007). Bicyclist injury severities in bicycle–motor vehicle accidents. *Accid. Anal. Prev., 39*, 238–251.

Klassen, J., El-Basyouny, K., & Islam, M. T. (2014). Analyzing the severity of bicycle-motor vehicle collision using spatial mixed logit models: a city of Edmonton case study. *Saf. Sci., 62*, 295–304.

Klop, J. R., & Khattak, A. J. (1999). Factors influencing bicycle crash severity on two-lane, undivided roadways in north carolina. *Transp. Res. Record, 1674*, 78–85.

Li, G., Braver, E. R., & Chen, L. H. (2003). Fragility versus excessive crash involvement as determinants of high death rates per vehicle-mile of travel among older drivers. *Accid. Anal. Prev., 35*, 227–235.

Li, Y., Ma, D., Zhu, M., Zeng, Z., & Wang, Y. (2018). Identification of significant factors in fatal-injury highway crashes using genetic algorithm and neural network. *Accid. Anal. Prev., 111*, 354–363.

Liu, J., Hainen, A., Li, X., Nie, Q., & Nambisan, S. (2019). Pedestrian injury severity in motor vehicle crashes: an integrated spatio-temporal modeling approach. *Accid. Anal. Prev., 132* 105272.

Liu, J., Khattak, A. J., Li, X., Nie, Q., & Ling, Z. (2020). Bicyclist injury severity in traffic crashes: a spatial approach for geo-referenced crash data to uncover non-stationary correlates. *J. Saf. Res.*.

Liu, X., Shen, L., & Huang, J. (1995). Analysis of bicycle accidents and recommended countermeasures in Beijing, China. *Transp. Res. Record, 1487*, 75–83.

Moore, D. N., Schneider, W. H., IV, Savolainen, P. T., & Farzaneh, M. (2011). Mixed logit analysis of bicyclist injury severity resulting from motor vehicle crashes at intersection and non-intersection locations. *Accid. Anal. Prev., 43*, 621–630.

Narkhede, S. (2018). Understanding auc-roc curve. *Towards Data Sci., 26*.

Prati, G., Pietrantoni, L., & Fraboni, F. (2017). Using data mining techniques to predict the severity of bicycle crashes. *Accid. Anal. Prev., 101*, 44–54.

Raihan, M. A., Alluri, P., Wu, W., & Gan, A. (2019). Estimation of bicycle crash modification factors (cmfs) on urban facilities using zero inflated negative binomial models. *Accid. Anal. Prev., 123*, 303–313.

Robartes, E., & Chen, T. D. (2017). The effect of crash characteristics on cyclist injuries: an analysis of virginia automobile-bicycle crash data. *Accid. Anal. Prev., 104*, 165–173.

Saha, D., Alluri, P., & Gan, A. (2015). Prioritizing highway safety manual's crash prediction variables using boosted regression trees. *Accid. Anal. Prev., 79*, 133–144.

Sivasankaran, S. K., & Balasubramanian, V. (2020). Exploring the severity of bicycle–vehicle crashes using latent class clustering approach in India. *J. Saf. Res., 72*, 127–138.

Stipancic, J., Zangenehpour, S., Miranda-Moreno, L., Saunier, N., & Granié, M. A. (2016). Investigating the gender differences on bicycle-vehicle conflicts at urban intersections using an ordered logit methodology. *Accid. Anal. Prev., 97*, 19–27.

VicRoads, 2013. Crashstats User Guide. http://data.vicroads.vic.gov.au/metadata/crashstats_user_guide_and_appendices.pdf. Online; accessed 04 Jun 2019. .

VicRoads, 2019. Crashes Last Five Years. https://vicroadsopendata-vicroadsmaps.opendata.arcgis.com/datasets/crashes-last-five-years. Online; accessed 22 Apr 2019. .

Rash-ha Wahi, R., Haworth, N., Debnath, A. K., & King, M. (2018). Influence of type of traffic control on injury severity in bicycle–motor vehicle crashes at intersections. *Transp. Res. Record, 2672*, 199–209.

Wall, S., Lee, D., Frangos, S., Sethi, M., Heyer, J., Ayoung-Chee, P., & DiMaggio, C. (2016). The effect of sharrows, painted bicycle lanes and physically protected paths on the severity of bicycle injuries caused by motor vehicles. *Safety, 2*, 26.
Wang, C., Lu, L., & Lu, J. (2015). Statistical analysis of bicyclists' injury severity at unsignalized intersections. *Traffic Injury Prev., 16*, 507–512.
Yan, X., Ma, M., Huang, H., Abdel-Aty, M., & Wu, C. (2011). Motor vehicle–bicycle crashes in beijing: irregular maneuvers, crash patterns, and injury severity. *Accid. Anal. Prev., 43*, 1751–1758.
Yasmin, S., & Eluru, N. (2018). A joint econometric framework for modeling crash counts by severity. *Transportmetr. A Transp. Sci., 14*, 230–255.
Zhang, Y., & Haghani, A. (2015). A gradient boosting method to improve travel time prediction. *Transp. Res. C Emerg. Technol., 58*, 308–324.
Zheng, Z., Lu, P., & Lantz, B. (2018). Commercial truck crash injury severity analysis using gradient boosting data mining model. *J. Saf. Res., 65*, 115–124.

Siying Zhu is a Ph.D. student in the School of Civil and Environmental Engineering, Nanyang Technological University, Singapore. She received the bachelor?s degree from Nanyang Technological University, Singapore, in 2017. Her research involves network modeling, urban infrastructure, data analysis, public transport, and transportation safety.

Journal of Safety Research 76 (2021) 314–326

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Assessing the likelihood of secondary crashes on freeways with Adaptive Signal Control System deployed on alternate routes

M Sabbir Salek [a,*], Weimin Jin [a], Sakib Mahmud Khan [a], Mashrur Chowdhury [a], Patrick Gerard [b], Saurabh Bikram Basnet [a], Mohammad Torkjazi [c], Nathan Huynh [c]

[a] Glenn Department of Civil Engineering, Clemson University, Clemson, SC 29634, USA
[b] Applied Statistics, School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC 29634, USA
[c] Department of Civil and Environmental Engineering, University of South Carolina, Columbia, SC 29208 USA

ARTICLE INFO

Article history:
Received 24 March 2020
Received in revised form 15 September 2020
Accepted 21 December 2020
Available online 12 January 2021

Keywords:
Secondary crash
Adaptive Signal Control Systems
Crash Likelihood Estimation
Logistic regression model
Unobserved heterogeneity

ABSTRACT

Introduction: Reducing the likelihood of freeway secondary crashes will provide significant safety, operational and environmental benefits. This paper presents a method for assessing the likelihood of freeway secondary crashes with Adaptive Signal Control Systems (ASCS) deployed on alternate routes that are typically used by diverted freeway traffic to avoid any delay or congestion due to a freeway primary crash. *Method:* The method includes four steps: (1) identification of secondary crashes, (2) verification of alternate routes, (3) assessment of the likelihood of secondary crashes for freeways with ASCS deployed on alternate routes and non-ASCS (i.e. pre-timed, semi- or fully-actuated) alternate routes, and (4) investigation of unobserved heterogeneity of the likelihood of freeway secondary crashes. Four freeway sections (i.e., two with ASCS deployed on alternate routes and two non-ASCS alternate routes) in South Carolina are considered. *Results and Conclusions:* Findings from the logistic regression modeling reveal significant reduction in the likelihood of secondary crashes for one freeway section (i.e., Charleston I-26 E) with ASCS deployed on alternate route. Other factors such as rear-end crash, dark or limited light, peak period, and annual average daily traffic contribute to the likelihood of freeway secondary crashes. Furthermore, random-parameter logistic regression model results for Charleston I-26 E reveal that unobserved heterogeneity of ASCS effect exists across the observations and ASCS are associated with the reduction of the likelihood of freeway secondary crashes for 84% of the observations (i.e., primary crashes). Location of the primary crash on the freeway is observed to affect the benefit of ASCS toward freeway secondary crash reduction as the primary crash's location determines how many upstream freeway vehicles will be able to take the alternate route. *Practical Applications:* Based on the findings, it is recommended that the South Carolina Department of Transportation (SCDOT) considers deploying ASCS on alternate routes parallel to freeway sections where high percentages of secondary crashes are found.

1. Introduction

Secondary crashes are defined as crashes that occur within the spatio-temporal impact range of a primary incident (Kitali et al., 2018) caused by traffic congestion due to occurrence of the primary incident. Although non-recurrent in nature, secondary crashes can impose substantial delay, emissions and operating costs, and they increase the crash risk for the vehicles upstream. Therefore, transportation agencies are seeking new technologies or policies of reducing possible secondary crashes on freeways. With more and more motorists having access to smart phone-based navigation tools such as Google Maps, Apple Maps, and Waze, drivers are more likely to divert to an alternate route for a reduced travel time when there is a major incident on the freeway. If such alternate routes are equipped with an advanced traffic management technology such as Adaptive Signal Control Systems (ASCS) that update signal parameters in real-time based on the fluctuating traffic demand, then the alternate routes with ASCS can better accommodate the diverted traffic from freeways as well as reduce the likelihood of secondary crashes on freeways, compared to the alternate routes without ASCS.

ASCS typically include algorithms that optimize and update the traffic signal parameters (i.e., cycle lengths, phase splits and

* Corresponding author.
E-mail addresses: msalek@clemson.edu (M.S. Salek), weiminj@clemson.edu (W. Jin), sakibk@g.clemson.edu (S.M. Khan), mac@clemson.edu (M. Chowdhury), pgerard@clemson.edu (P. Gerard), sbasnet@clemson.edu (S.B. Basnet), nathan.huynh@sc.edu (N. Huynh).

https://doi.org/10.1016/j.jsr.2020.12.020
0022-4375/

sequences, and offsets) in real-time (Lowrie, 1990; Mirchandani & Head, 2001; Gartner et al., 2002; Jin et al., 2019; Jin et al., 2020; Jin et al., 2021). Operational benefits of ASCS are significant in both corridor and network operations as reported in the literature (So et al., 2014; Khattak, 2016; Elkins & Niehus, 2012; Fontaine et al., 2015; Kergaye et al., 2009; Eghtedari, 2006). Prior to the deployment of ASCS, conventional signal systems based on time of day (TOD) were used in many locations where signal plans are typically pre-set and adjusted every 2-3 years. These conventional signal systems are unable of handling highly variable traffic demand. With real-time traffic signal parameters adjustment capability, ASCS are better at handling highly variable traffic demand or traffic congestions caused by traffic incidents or special events. Thus, any alternate route with ASCS can be used to potentially improve the traffic operations on a freeway in case any incident happens on that freeway. However, to the best of our knowledge there exists no such study in the literature that evaluated the possible benefits of having an ASCS deployed on alternate route to a freeway section toward the reduction of freeway secondary crashes.

In this paper, the authors present a method for assessing the likelihood of secondary crashes on freeways with alternate routes where ASCS have been deployed. The applicability of the method is firstly demonstrated with two freeway sections where ASCS are deployed in several intersections within the alternate routes used by diverted freeway traffic when a major crash occurs on the freeway sections. The authors develop fixed-parameter binary logistic regression models for Charleston I-26 (Eastbound and Westbound) in South Carolina to investigate if the presence of an ASCS deployed alternate route is associated with the reduction of the likelihood of freeway secondary crashes. This study also develops fixed-parameter binary logistic regression models for two freeway sections with non-ASCS (i.e., pre-timed, semi- or fully-actuated) alternate routes to examine if the likelihood of secondary crashes differs between freeways with ASCS deployed on alternate routes and freeways with non-ASCS alternate routes. In addition, the effect of ASCS on the likelihood of secondary crashes on a freeway may vary across observations (i.e., primary crashes). To capture unknown variations in the effect of ASCS across the observations (which the authors refer to as "unobserved heterogeneity" in this paper) the authors develop a random-parameter binary logistic regression model to account for observation-specific variations in the effects of ASCS and to provide more accurate inferences. Note that there is a significant number of studies in the literature on the likelihood estimation of secondary crash occurrences using logistic regression models (Kitali et al., 2019; Karlaftis et al., 1999; Goodall, 2017; Xu et al., 2016; Yang et al., 2014). However, these studies did not consider the effect of having an advanced traffic management or control technology, such as ASCS, deployed on alternate route on freeway secondary crashes. Thus, the primary contribution of this paper lies in exploring the safety benefit of having an ASCS deployed alternate route toward the reduction of the likelihood of freeway secondary crashes.

The remainder of the paper is organized as follows: literature review, method, data sources and variables, analyses and results, conclusions, and practical applications. The literature review section reviews the identification of secondary crash and the likelihood estimation modeling of secondary crashes. The method section focuses on the procedure of identification of secondary crashes, verification of alternate routes, and modeling the likelihood of secondary crashes. The data sources and variables section explains the crash data received from the South Carolina Department of Transportation (SCDOT) and variables that are used in the models. The analyses and results section presents the model estimation results, and conclusions summarize the main findings of this paper. Practical applications section presents policy implications.

2. Literature review

In this section, the authors review the previous work related to the likelihood estimation of secondary crashes focusing on the secondary crash identification methods and the likelihood estimation models.

2.1. Identification of secondary crashes

Several studies have sought to investigate the criteria to identify secondary crashes. Table 1 shows a summary of these criteria. Raub (1997a, 1997b)) considered any crashes within the time period of a primary crash plus 15 minutes and within a mile from the primary crash in the upstream as secondary crashes. Based on these criteria, the author identified secondary crashes and found that 81 primary crashes were followed by 97 secondary crashes. The author concluded that 1 of every 11 incidents that occurred in Rolling Meadows (between January 9, 1995 and February 5, 1995) was associated with one or more secondary crashes. Karlaftis et al. (1999) analyzed 5 years of incident data on Borman Expressway in Illinois to identify the primary crash characteristics that led to the secondary crashes. Latoski et al. (1999) analyzed the data from portions of I-80, I-94 and I-65 in North West Indiana. Both Karlaftis et al. (1999) and Latoski et al. (1999) considered 3 miles upstream of the primary crash and the clearance time plus 15 minutes following the primary crashes to identify secondary crashes.

Moore et al. (2004) studied 84,684 crashes in California. The authors considered crashes that occurred within a 2-hour time period and 2 miles in both directions of the primary crashes as an identification measure for the secondary crashes. Hirunyanitiwattana and Mattingly (2006) studied the characteristics of secondary crashes using two years of crash data from the California highway system from 1999 to 2000. They found that a secondary crash is the one that occurred within an hour and 2 miles upstream of the primary crash. They also found that the proportion of secondary crashes was higher in urban areas compared to rural areas. Yang et al. (2013) developed a method based on binary speed contour plot to account for the dynamic characteristics of spatio-temporal impact range in identifying secondary crashes. This study used sensor data from a 27-mile urban highway section in New Jersey. The authors found that almost 50% of the secondary crashes occurred within a 2-mile range in the upstream, and 75% of the secondary crashes occurred within up to 2 hours of the primary crashes.

2.2. Modeling of the likelihood of secondary crashes

Karlaftis et al. (1999) primarily investigated contributing factors that induce a secondary crash. They developed a fixed-parameter binary logistic regression model using the attributes of the primary crashes to estimate the possibility of a secondary crash. The study

Table 1
Review of Secondary Crash Identification Criteria.

Author(s), year	Secondary crash identification criteria	Road type
Raub, 1997a,b	Clearance time + 15 minutes, 1 mile	Urban arterial
Karlaftis et al., 1999	Clearance time + 15 minutes, 1 mile	Freeway/expressway
Latoski et al., 1999	Clearance time + 15 minutes, 3 miles	Freeway
Moore et al., 2004	2 hours, 2 miles	Freeway
Hirunyanitiwattana & Mattingly, 2006	1 hour, 2 miles	Urban/rural freeway/highway
Yang et al., 2013	2 hours, 2 miles	Urban highway

concluded that the attributes such as clearance time, season, type of vehicle involved, and lateral location of the primary crash were the most significant factors for increased likelihood of secondary crashes. Goodall (2017) developed a fixed-parameter binary logistic regression model to predict the occurrence of secondary crashes over time. Three contributing factors were considered, namely, (a) whether congestion occurred or did not occur due to an incident, (b) the incident duration, and (c) the approximate number of vehicles that encountered an incident (if no congestion) or queue due to an incident (if congestion was present) in the same direction. The results revealed that, for every 2–3 minutes spent for a congested scenario, the secondary crash occurrence probability increases approximately by 1%. Xu et al. (2016) used the Bayesian random effect logit model to develop a secondary crash risk prediction model. The model associated the prediction probability of secondary crashes with variables such as real-time traffic variables (e.g., average speed, traffic volume, standard deviation of detector occupancy), primary crash characteristics (e.g., date and time of primary crash, primary crash severity and crash type), weather conditions, and geometric characteristics. The most significant real-time traffic variables were traffic volume, average speed, the standard deviation of detector occupancy, and volume difference between adjacent lanes. The study concluded that the secondary crash prediction accuracy can be increased by 16.6% by including traffic flow variables. Yang et al. (2014) assessed the risk of secondary crashes on major highways. A rare event logistic regression model was used to study the secondary crashes. The authors concluded that one additional minute increase in the incident duration could increase the likelihood of secondary crashes by 1.2%.

The fixed-parameter binary logistic regression model cannot capture unobserved heterogeneity since the model assumes global effect of each predictor for all observations. Therefore, a random-parameter binary logistic regression model is deployed in this paper to study the unobserved heterogeneous effect of the predictors on the likelihood of freeway secondary crashes across observations.

3. Method

The method to assess the likelihood of secondary crashes on freeways with alternate routes includes four steps:

(1) Identification of secondary crashes (using fixed spatio-temporal criteria and other factors such as manner and probable cause of collision)
(2) Verification of alternate routes (with SCDOT and travel time data from the alternate routes)
(3) Modeling the likelihood of secondary crashes for both ASCS deployed on alternate routes and non-ASCS alternate routes using fixed-parameter binary logistic regression models
(4) Investigation of unobserved heterogeneity of ASCS using a random-parameter binary logistic regression model.

This section of the paper explains these four steps of the method in details.

3.1. Identification of secondary crashes

Selecting a spatio-temporal criterion for identifying the secondary crashes on freeways is a challenging task. Secondary crashes are typically induced by primary crashes that cause adverse effects on the traffic flow. These impacts of primary crashes on the traffic flow vary depending on many factors such as the number of blocked lanes, clearance time, and crash severity. However, lane blockage and clearance time information are not available for the study corridors. Thus, the individual impact of each primary crash cannot be determined. Therefore, the authors consider a fixed spatio-temporal range as a primary criterion for the identification of secondary crashes. The authors consider a crash to have the possibility of being induced by a primary crash if it occurs within a 1-hour period after the primary crash and within a 2-mile range in the upstream of the primary crash. The authors use this fixed spatio-temporal range for the secondary crash identification as there is no available real-time traffic volume data for the four freeway sections considered in this paper that may be used to develop dynamic ranges for the secondary crash identification.

The temporal threshold used in this paper to identify the secondary crashes (i.e., up to 1-hour after a crash occurrence on the freeway) can be justified based on the overall crash detection, response and clearance time in South Carolina. According to Chowdhury et al. (2007), the incident detection time of Traffic Management Center (TMC) is 1–5 minutes for South Carolina and arrival of the first responder takes 9–10 minutes after that. In addition, according to the SCDOT State Highway Emergency Program's (SCDOT SHEP, 2020) 2019 database (obtained from SCDOT), the average clearance time for Charleston is 38.5 minutes and for Richland-Lexington is 38.7 minutes. Therefore, on average we get 48–54 minutes by combining the detection, response, and clearance time that is lower than the selected temporal threshold of 1 hour for the secondary crash identification. Using this 1-hour temporal threshold, the authors can decide on the spatial threshold for the secondary crash identification by observing the relative change in the number of identified secondary crashes as we vary the spatial threshold. Fig. 1 presents the relative change in the number of identified secondary crashes for spatial threshold ranging from 1 mile to 4 miles. As observed from Fig. 1, an increment of the spatial threshold from 2 miles to 2.5 miles causes less than 10% relative change (i.e., less than 10% relative increase) in the number of identified secondary crashes. While increasing the spatial threshold beyond 3 miles can cause this relative change to be lower than 5%, the authors choose to use a fixed spatial threshold of 2 miles as it makes more sense with the SCDOT practitioners as an upper limit for the spatial impact range based on their experience.

The limitation of using a fixed spatio-temporal criterion is that it may not capture all the secondary crashes since the impact of

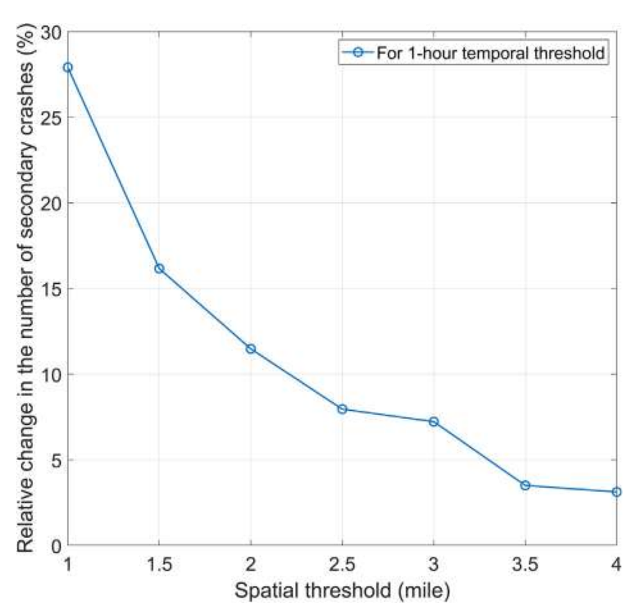


Fig. 1. Relative change in the number of secondary crashes with varying spatial threshold.

some primary crashes on traffic may exceed the predefined spatial and temporal thresholds. However, fixed spatial and temporal thresholds may include some crashes that may have been caused by some other factors. Therefore, the authors utilize two criteria, namely, "manner of collision" and "probable cause of collision," after applying the spatial and temporal thresholds for filtering the secondary crashes. The rationale behind using these two criteria is to prevent any crashes from being misclassified as secondary crashes that occurred due to some other reasons. To be more specific, if the "manner of collision" for a crash that falls within the spatio-temporal impact range of a primary crash is listed as a "head-on collision," then the authors do not label that crash as a secondary crash. Typically, a head-on collision can occur only between two vehicles traveling in opposing directions. Therefore, crash data related to either one of the two vehicles involved in a head-on collision should not be labeled as a secondary crash. Apart from that, the other types of crashes, for example, rear-end, angle, and side-swipe crashes, are not discarded this way because it is not reasonable to assume that these types of crashes cannot be caused by primary crash's impact. Similarly, if the "probable cause of collision" for a crash that falls within the spatio-temporal impact range of a primary crash is listed as either one of (a) tire/wheel failure, (b) mechanical failure of the vehicle, (c) debris/obstruction or animal on the roadway, and (d) medical related, then the authors deem it reasonable to not label that crash as a secondary crash. Thus, the authors do not label a crash within the spatio-temporal impact range of a primary crash as a secondary crash, if it is not reasonable to be labeled as a secondary crash. While this information might be subject to police misspecification and reporting practice, the authors could not find any additional means to cross-validate this information. However, the authors observe that after satisfying the spatio-temporal criterion, only a few crashes are not considered as secondary crashes due to their "manner of collision" or "probable cause of collision." For example, for Charleston I-26, only 6 out total 3562 crashes, and for Richland-Lexington I-26, no crashes are discarded as secondary crashes based on the manner of collision or probable cause of collision. Fig. 2 presents this crash identification procedure with a flow diagram.

3.2. Verification of alternate routes

The authors investigate the parallel arterials of the freeway sections for alternate route verification. Firstly, the authors verify the alternate routes with SCDOT. Then, the authors utilize real-time travel data to investigate the changing traffic conditions of the parallel arterials in the event of crashes on the freeways. Hourly travel time data recorded by ClearGuide (SCDOT Iteris ClearGuide, 2020) is used to observe how the average travel time of the parallel arterial changes in the 1-hour after period when a crash occurs on the freeway section. For each crash on the freeway, weighted average travel time in 1-hour after period is computed from the hourly travel time data. For example, if a crash occurs at 05:25 PM, then the weighted average travel time on the ASCS-deployed alternate route for the 1-hour after period (05:25 PM to 06:25 PM) is calculated as follows,

Weighted average travel time (05:25 PM to 06:25 PM) = $\frac{35}{60}\times$ (average travel time from 05:00 PM to 06:00 PM) + $\frac{25}{60}\times$ (average travel time from 06:00 PM to 07:00 PM)

The weighted average travel time is then compared with the historical weighted average of travel time for that period of time of the day. Hourly travel time data recorded by ClearGuide (SCDOT Iteris ClearGuide, 2020) from four consecutive months around the time when the crash occurred is used to compute historical weighted average of hourly travel time data. Historical weighted average of travel time for 1-hour after period of the crash is then computed from the historical average of hourly travel time data similarly as shown in the last example above. The historical weighted average is computed separately for weekdays and weekends. The weighted average travel time for 1-hour after period of a crash occurrence is compared with the 95th percentile of historical weighted average of travel time for that time period. If the weighted average exceeds the 95th percentile of the historical weighted average of travel time, then the change in travel time (through the alternate route) due to the crash occurrence on the freeway is considered to be significant.

3.3. Fixed-parameter binary logistic regression model

The authors develop a fixed-parameter binary logistic regression model to evaluate the likelihood of secondary crashes, as the secondary crash occurrence is a binary outcome (occurrence or non-occurrence) that can depend on many factors.

A fixed-parameter binary logistic regression (logit) model can be developed to evaluate the likelihood of a secondary crash occurrence, which is formulated as follows,

$$\log\left(\frac{P(y=1|\boldsymbol{X})}{1-P(y=1|\boldsymbol{X})}\right)=\boldsymbol{X\beta} \quad (1)$$

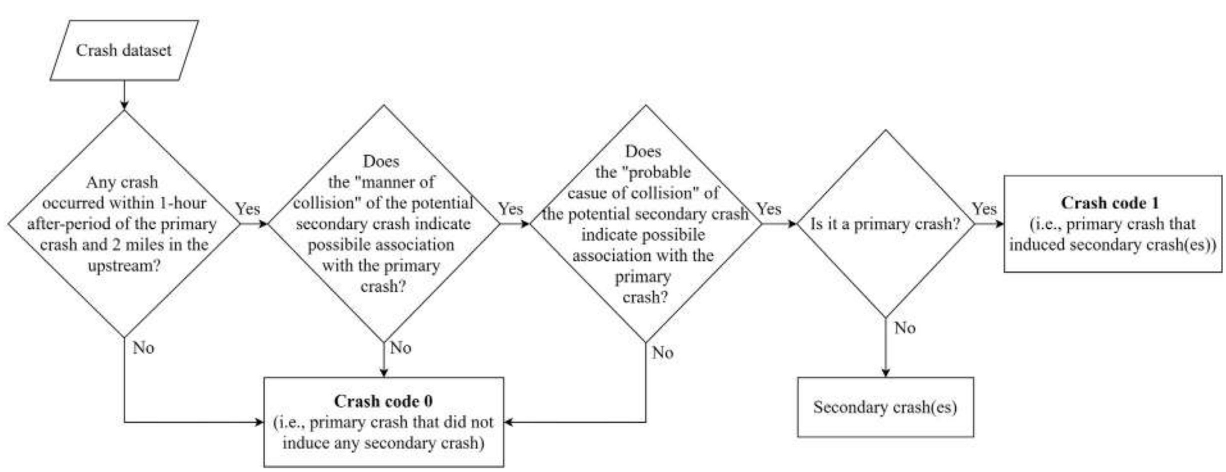


Fig. 2. Crash identification procedure.

where, $P(y = 1|\boldsymbol{X})$= the conditional probability of a secondary crash occurrence given a primary crash occurred

X = the vector of explanatory variables associated with the primary crashes

$\boldsymbol{\beta}$=the vector of coefficients corresponding to the explanatory variables

Equation (1) presents a general form of the fixed-parameter binary logistic regression model. The authors develop corridor-specific fixed-parameter binary logistic regression models for both freeways with ASCS deployed on alternate routes and freeways with non-ASCS alternate routes from equation (1) and identify statistically significant explanatory variables. The response variable, y in equation (1) is equal to 1 if a secondary crash occurred or 0 otherwise.

The authors use open-source R software to perform the regression analyses. The generalized linear model function "glm" in R is used to estimate the coefficients of the logistic regression model. The Iterative Weighted Least Squares (IWLS) is used in this function to obtain the Maximum Likelihood Estimation (MLE) of $\boldsymbol{\beta}$. Using the fitted model, the effect of the k-th explanatory variable on the occurrence of a secondary crash can be evaluated by Odds Ratios (ORs) given by,

$$OR = e^{\beta_k} \tag{2}$$

where, β_k is the coefficient of the k-th explanatory variable in the fitted model.

The Variance Inflation Factor (VIF) is used to check for potential Multi Collinearity (MC). Many researchers used a VIF of 10 to indicate excessive or severe MC issue (O'Brien, 2007). Akaike Information Criteria (AIC) is compared among different candidate models and the model with the lowest AIC value is preferred.

3.4. Random-parameter binary logistic regression model

Compared to the fixed-parameter binary logistic regression model, a random-parameter binary logistic regression (logit) model can capture unobserved heterogeneity across observations. Equation (1) can be rewritten as,

$$\log\left(\frac{P(y_i = 1|\boldsymbol{X}_i)}{1 - P(y_i = 1|\boldsymbol{X}_i)}\right) = \boldsymbol{X}_i\boldsymbol{\beta}_i;\ i = 1, 2, 3, \ldots..n \tag{3}$$

$$\boldsymbol{\beta}_i \sim g(\boldsymbol{\beta}_i|\boldsymbol{\theta}) \tag{4}$$

where, $\boldsymbol{X}_i$ is a vector of the explanatory variables of observation (i.e., primary crash) i, and $\boldsymbol{\beta}_i$ is a vector of the coefficients.

In the random-parameter binary logit model, $\boldsymbol{\beta}_i$ is allowed to be varying for each observation i rather than fixed for all observations. The distribution, $g(\boldsymbol{\beta}_i|\boldsymbol{\theta})$ is specified to enable $\boldsymbol{\beta}_i$ vary across each observation, where $\boldsymbol{\theta}$ is a vector including the mean and variance of a random distribution.

$\boldsymbol{\beta}_i$ can be written as $\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{L}\boldsymbol{\omega}_i$, where $\boldsymbol{\beta}$ is the vector of the mean of coefficients. Note that, each coefficient can be written as $\beta_{ki} = \beta_k + \sigma_k\omega_i$. β_{ki} is k-th element in $\boldsymbol{\beta}_i$. $\boldsymbol{\omega}_i$ is a vector of random variables that follow random distributions. $\boldsymbol{L}$ is a diagonal matrix that contains the standard deviations of the coefficients, σ_k. In this paper, β_{ki} is considered to follow a normal distribution, which is specified as $\beta_{ki} \sim N(\beta_k, \sigma_k^2)$. The normal distribution specification for β_{ki} provides a better model fit, compared to other possible distributions such as log normal distribution based on our analysis.

To explore the unobserved heterogeneity of the parameters (i.e., coefficients in the model) across observations, the conditional mean of the parameters is estimated. The estimator of the conditional mean of the random parameters (Sarrias, 2016) is obtained by the Simulated Maximum likelihood (SML) procedure, which is expressed as:

$$\hat{E}(\boldsymbol{\beta}_i|data_i) = \sum_{r=1}^{R}\left(\frac{\hat{P}(y_i|\boldsymbol{X}_i, \boldsymbol{\beta}_{ir})}{\sum_{r=1}^{R}\hat{P}(y_i|\boldsymbol{X}_i, \boldsymbol{\beta}_{ir})}\right)\hat{\boldsymbol{\beta}}_{ir} \tag{5}$$

where, $\hat{\boldsymbol{\beta}}_{ir} = \hat{\boldsymbol{\beta}} + \hat{\boldsymbol{L}}\boldsymbol{\omega}_{ir}$; y_i is the response variable (1 if a secondary crash occurred; 0 otherwise). $data_i$ stands for the explanatory variables associated with each observation. $\hat{P}(\boldsymbol{y}_i|\boldsymbol{X}_i, \boldsymbol{\beta}_{ir})$ is the estimated simulated probability for the observation i evaluated at the r-th random draw of $\boldsymbol{\beta}_i$; R is the total number of random draws in the SML procedure. In the estimation of conditional mean of the parameters, a Halton random number generator with a standard uniform distribution, $U(0, 1)$ generates the random draws. Detail Halton draws procedure can be found in Sarrias (2016).

The SML procedure is conducted to obtain the model estimation results using the "Rchoice" library in the R software. 100 Halton draws are used in the SML procedure for the purpose of model estimation (Sarrias, 2016). Likelihood ratio test is used to compare the performance between the random-parameter models and the fixed-parameter models (Washington et al., 2020).

4. Data sources and variables

The authors consider two freeway sections with ASCS deployed on alternate routes to investigate the likelihood of secondary crash occurrences (see Fig. 3(a)): (1) an 8.92-mile section of Charleston I-26 E, (2) a 9.6-mile section of Charleston I-26 W. These two corridors are referred as "Freeways with ASCS deployed on alternate routes" in rest of the paper as they have an ASCS deployed parallel arterial (i.e., US 52). There are three main components in the ASCS deployed on US 52: the vehicle detection, local traffic controller(s), and central server. The central server processes data and generates optimized signal timings. The local traffic controller collects vehicle detection data and receives the signal timings from the central server. The primary objective of the ASCS algorithm is to minimize total traffic delays of the intersections while ensuring suitable progression bandwidth of the main corridors. For the Charleston I-26 freeway sections, ASCS was deployed at 17 intersections of parallel US-52 in October 2016. US 52 is considered as an alternate route (verified by SCDOT and with ClearGuide data) for the diverting traffic of I-26 Eastbound and Westbound sections in the event of a freeway primary crash. The functional class of US 52 is principal arterial.

This study also uses freeways with non-ASCS (i.e., pre-timed, semi- or fully-actuated) alternate routes similar to the freeways with ASCS deployed on alternate routes to examine if the effect of the after-period indicator (i.e., an explanatory variable used in our regression models which will be explained in the next section) differs between these freeways. The authors select two freeway sections with non-ASCS alternate routes that have comparable lengths, annual average daily traffic (AADT) of the freeway sections and the same functional classes as the freeways with ASCS deployed on alternate routes (see Fig. 3(b)). The freeways with non-ASCS alternate routes are: (1) a 7.75-mile section of Richland-Lexington I-26 E, (2) a 7.64-mile section of Richland-Lexington I-26 W. In addition, both the ASCS-deployed sites and the non-ASCS sites considered for this study are located within the jurisdiction of the same Department of Transportation in the United States (i.e., South Carolina Department of Transportation or SCDOT). Therefore, the authors assume similar management and maintenance characteristics, such as pavement maintenance, traffic management, and enforcement for the corridors considered in this study. A comparison of characteristics of the freeways with ASCS deployed on alternate routes and freeways with non-ASCS alternate routes are presented in Table 2.

For the analysis, the authors use "crash code" (i.e., crash code 0 and crash code 1 to indicate primary crashes that did not induce

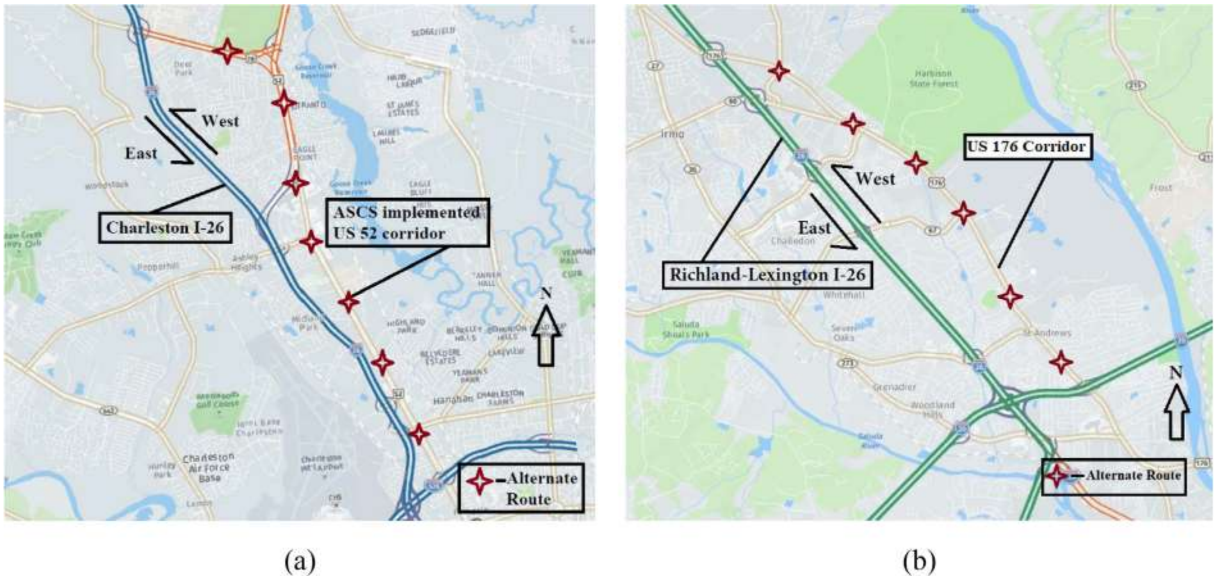


Fig. 3. (a) Charleston I-26 with ASCS deployed on alternate route US 52, and (b) Richland-Lexington I-26 with non-ASCS alternate route US 176.

Table 2
Comparison of the freeways with ASCS deployed on alternate routes and non-ASCS alternate routes.

Corridor name	Corridor length	Mean AADT	Alternate route	Functional class of the alternate route
Freeways with ASCS deployed on alternate routes				
Charleston I-26 E	8.92 miles	149,852	US 52	Principal arterial
Charleston I-26 W	9.6 miles	145,875	US 52	Principal arterial
Freeways with non-ASCS alternate routes				
Richland-Lexington I-26 E	7.75 miles	119,699	US 176	Principal arterial
Richland-Lexington I-26 W	7.64 miles	115,983	US 176	Principal arterial

any secondary crashes and primary crashes that induced one or more secondary crashes, respectively) as the response variable. Tables 3 and 4 present a summary of the primary crashes in the crash data used for analysis based on the crash codes.

To evaluate the effect of ASCS deployment on the likelihood of freeway secondary crashes, the authors extract a total of 52 months crash data for Charleston I-26 (East and West) corridors, i.e., 26 months (September 2014 to October 2016) for the before period of the ASCS deployment, and 26 months (November 2016 to December 2018) for the after-period of the ASCS deployment. The same period of data is extracted for Richland-Lexington I-26 (East and West) corridors. The complete extracted dataset consisting of both primary and secondary crashes includes a total of 1,757 crashes on the Charleston I-26 E section (772 crashes in the before-period and 985 crashes in the after-period), and a total of 1805 crashes on the Charleston I-26 W section (849 crashes in the before-period and 956 crashes in the after-period). Note that the complete extracted dataset includes both primary and secondary crashes. After identifying the primary crashes with "Crash code 0" and "Crash code 1", the authors conducted further analyses with the primary crashes only. The frequencies and percentages of these primary crashes are presented in Table 3.

Table 3
Summary of response variables of the freeways with ASCS deployed on alternate routes.

Corridor name	Crash code	Frequency	Percentage
Charleston I-26 E	0	1443	91.04%
	1	142	8.96%
Charleston I-26 W	0	1518	92.17%
	1	129	7.83%

Table 4
Summary of response variables of the freeways with non-ASCS alternate routes.

Corridor name	Crash code	Frequency	Percentage
Richland-Lexington I-26 E	0	1233	90.73%
	1	126	9.27%
Richland-Lexington I-26 W	0	1368	90.66%
	1	141	9.34%

For each freeway section, the crash data does not only include crashes that occurred on the freeway but also includes the crashes that occurred on the entrance ramps to the freeway section. The rationale for including the entrance ramp crashes in the analyses is that when a crash occurs on the freeway, sometimes if it creates a major congestion, then traffic can back up to the entrance ramp and can cause secondary crashes on the ramps as well. Therefore, the entrance ramp crashes are used in the dataset only to check if there was any secondary crash on the ramps induced by a primary crash on the freeway. Table 5 lists the number of entrance ramps included in the crash data set for the secondary crash detection.

SCDOT provides the crash data for the analyses presented in this paper. The crash data includes several attributes such as collision time, AADT, light condition, roadway surface condition, manner of collision, weather condition, first harmful event, and probable cause of crash. Variables such as light condition (dawn, daylight, dusk or dark), roadway surface condition (dry, icy, wet or snowy), weather condition (adverse or not adverse condition), and manner of collision (rear-end, angle, head-on, side-swipe, etc.) help to account for various possible attributes that may have effects on the secondary crash occurrence. The authors do not use real-time traffic volume data since there is a lot of missing data in the whole study period.

Table 5
Number of entrance ramps included in the crash data set.

			Number of entrance ramps
Freeways with ASCS deployed on alternate routes	Charleston I-26	Eastbound Westbound	11 15
Freeways with non-ASCS alternate routes	Richland-Lexington I-26	Eastbound Westbound	8 8

"Peak period" is used as a predictor for modeling the likelihood of secondary crash occurrences on all the freeway sections. For each freeway section, the authors analyze the hourly average travel time recorded by ClearGuide (SCDOT Iteris ClearGuide, 2020) to define corridor-specific peak periods. The authors observe hourly average travel time for weekdays and weekends separately. However, no significant weekend peak periods are detected. Peak periods only exist on weekdays for the freeway sections considered here. Table 6 presents the corridor-specific weekday peak periods that are considered for logistic regression modeling. Note that, both Charleston I-26 E and Richland-Lexington I-26 E experience AM peak periods as traffic goes in to the center of the cities during this time (Charleston I-26 E goes in to the center of Charleston, and Richland-Lexington I-26 E goes in to the center of Columbia). Similarly, Charleston I-26 W and Richland-Lexington I-26 W experience PM peak periods as traffic comes out of the center of the cities during this time.

"After-period indicator for freeways with ASCS deployed on alternate routes" is used as a predictor to investigate the effect of ASCS deployment in the alternate route on the likelihood of freeway secondary crashes. In the models, "after-period indicator for freeways with ASCS deployed on alternate routes" is specified as 1 if a crash occurs in the after-period of ASCS deployment, and 0 if a crash occurs in the before-period of ASCS deployment. For the freeways with non-ASCS alternate routes, ASCS was not deployed on the alternate routes. The authors still include a predictor called "after-period indicator for freeways with non-ASCS alternate routes" with same temporal division as the freeways with ASCS deployed on alternate routes in order to examine if the effect of the temporal division (before-after period) differs between the freeways with ASCS deployed on alternate routes and the freeways with non-ASCS alternate routes.

"Temporal trend" variable is included to account for long-term temporal trends in safety due to unobserved factors such as long-term roadway conditions, weather conditions, and improvements in vehicular technologies (Persaud et al., 2010). The "Temporal trend" variable is coded as numerical values. For example, if a crash occurs in 2014, it is specified as 0, if a crash occurs in 2015, it is specified as 1, and so on.

Table 7 summarizes all the variables that are considered for the analysis of the likelihood of freeway secondary crashes.

Table 6
Corridor-specific weekday peak periods.

Corridor type	Corridor name	Corridor-specific weekday peak period
Freeways with ASCS deployed on alternate routes	Charleston I-26 E	5:30 AM to 8:30 AM
	Charleston I-26 W	3.00 PM to 6.00 PM
Freeways with non-ASCS alternate routes	Richland-Lexington I-26 E	6.30 AM to 8.30 AM
	Richland-Lexington I-26 W	3.30 PM to 6.30 PM

Table 7
Model Variables.

Category	Variable name	Description
Response variable	Crash code	1 – Primary crash that induces secondary crash(es) 0 – Primary crash that does not induce any secondary crash
Explanatory variables	After-period indicator	1 – Crash occurs in the after-period of ASCS deployment 0 – Crash occurs in the before-period of ASCS deployment
	Light condition	1 – Dawn, dusk, dark or limited light 0 – Daylight
	Roadway surface condition	1 – Icy, snowy or wet 0 – Dry
	Weather condition	1 – Adverse weather 0 – otherwise
	Rear end	1 – Primary crash is rear end 0 – otherwise
	Angle crash	1 – Primary crash is angle crash 0 – otherwise
	Weekday	1 – Primary crash occurs on a weekday 0 – otherwise
	Peak period	1 – Primary crash occurs during peak period 0 – otherwise
	Crash severity	Numerical values indicating five levels (0, 1, 2, 3, and 4); 0 = no injury, 1 = possible injury, 2 = non-incapacitating injury, 3 = incapacitating injury, 4 = fatal
	Temporal trend	Numerical values. 0 if a crash occurs in 2014, 1 if a crash occurs in 2015, 2 if a crash occurs in 2016, 3 if a crash occurs in 2017, 4 if a crash occurs in 2018.
	AADT	Numerical values. log(AADT) is used for scaling down purpose.

5. Analyses and results

5.1. Verification of alternate routes

Table 8 presents the results from the alternate routes verification for all the corridors. It is observed that for over 40% of the crashes occurred on the freeway sections from September 1, 2018 to December 31, 2018 (as ClearGuide travel time data are not available before September 1, 2018 and the study period for all the corridors ends on December 31, 2018), travel time increased significantly compared to the 95th percentile of the historical weighted average travel time in the corresponding parallel alternate routes indicating that drivers often use these routes when a crash occurs on the freeway.

5.2. Fixed-parameter binary logistic regression model results

Based on the explanatory variables considered here, equation (1) can be rewritten as follows,

$$\log\left(\frac{p(y=1|X)}{1-p(y=1|X)}\right)=\beta_0+\beta_1 \times(\text{after-period indicator of ASCS deployment}) +\beta_2\times(\text{light condition})+\beta_3 \times(\text{roadway surface condition})+\beta_4 \times(\text{weather condition})+\beta_5 \times(\text{rear-end crash})+\beta_6\times(\text{angle crash}) +\beta_7\times(\text{weekday})+\beta_8\times(\text{peak period}) +\beta_9\times(\text{crash severity})+\beta_{10} \times(\text{temporal trend})+\beta_{11}\times\log(\text{AADT}) \tag{6}$$

Table 8
Verification of the alternate routes with travel time information.

	Study corridor name	% of crashes on the freeway that caused average travel time to increase significantly in the alternate route
Freeways with ASCS deployed on alternate routes	Charleston I-26 E	44.76%
	Charleston I-26 W	51.56%
Freeways with non-ASCS alternate routes	Richland-Lexington I-26 E	47.42%
	Richland-Lexington I-26 W	42.02%

The authors apply the fixed-parameter binary logistic regression model (as presented in equation (6)) explained in the research method section to the crash datasets for all the study corridors. Table 9 presents the model estimation results of "After-period indicator variable of ASCS deployment" for the freeways with ASCS deployed on alternate routes and freeways with non-ASCS routes (based on corridor-specific fixed-parameter binary logistic regression models). For all the corridor-specific models, the authors check for any existing multicollinearity using the VIF. For all the ASCS and the non-ASCS corridors, the maximum VIF is found to be less than or equal to 3.5. Therefore, it is assumed the multicollinearity does not exist among the explanatory variables as VIF < 10.

As explained before, the Odds Ratio (OR) is defined as the ratio of the odds of an outcome occurring by exposure of a variable to the odds of the outcome occurring in the absence of that exposure. From the odds ratios, percentage changes in the secondary crash occurrence odds are evaluated. In Table 9, the odds ratios and percentage changes in the secondary crash occurrence odds are displayed only if the predictor (i.e., after-period indicator of ASCS deployment) is found to be statistically significant at a 0.1 significance level.

As shown in Table 9, for Charleston I-26 E, a 47.32% reduction of the likelihood of secondary crashes is associated with the after period of ASCS deployment. However, for Charleston I-26 W, the fixed-parameter binary logistic regression model cannot reveal any statistical significance of ASCS deployment for reducing the likelihood of secondary crashes. As the same parallel arterial may not be always convenient as an alternate route for the drivers traveling in opposing directions, the effect of ASCS deployment on the alternate routes can also be different for opposing traffics on the freeway. Also, opting for an alternate route depends on many factors such as drivers' behavior, freeway crash severity, lane blockage, and number of vehicles involved in the crash.

As mentioned in the research method section, for freeways with non-ASCS alternate routes, "after period indicator for freeways with alternate non-ASCS corridors" is included in the model as a predictor in order to observe if the temporal division used for ASCS deployment's before and after period signifies anything. In Table 9, it is observed that this variable is not significant for the freeways with non-ASCS alternate routes. It indicates that, this temporal division of the study period signifies nothing for the freeways with non-ASCS alternate routes as there was no ASCS deployment.

Table 10 presents the other statistically significant predictors and their corresponding coefficients estimate from logistic regression model. For both freeways with ASCS deployed on alternate routes and non-ASCS alternate routes, frequently observed statistically significant variables include- rear-end crashes, light condition, weekday, and AADT. Note that, crash severity and roadway surface condition are not included in Table 10, as for none of the corridors these two variables are significant.

The authors perform additional analysis for Charleston I-26 E with traffic count and speed data collected from the SCDOT Traffic Polling and Analysis System (SCDOT Traffic Counts, 2020). The authors perform this analysis exclusively for Charleston I-26 E to prove that the favorable effect of ASCS found for Charleston I-26 E is not a contribution of reduced crash exposure (i.e., lower freeway traffic counts after a crash occurrence) or reduced speed on the freeway. The authors compute the weighted average traffic count and speed on the freeway for the 1-hour after period of each crash occurrence based on the hourly traffic count and speed data collected from SCDOT Traffic Polling and Analysis System (SCDOT Traffic Counts, 2020). The weighted average of traffic count and speed is computed in the same way as the weighted average travel time computation explained in subsection "3.2 Verification of Alternate Routes" of section "3 METHOD."

In Fig. 4, the authors present two bar charts that present the percentages of "Crash code 1" (i.e., crashes that induced secondary crashes) across various ranges of traffic counts and speed on the freeway during the 1-hour after period of a freeway crash occurrence. For each range shown in Fig. 4(a) and (b), two separate bars are used to show the percentages of "Crash code 1" during the before-period and the after-period of ASCS deployment. In Fig. 4 (a), the authors combine vehicle counts during the 1-hour after period of freeway crash occurrences ranging from 0 to 3000 vehicles/hour into one bar and from 4000 to 6000 vehicles/hour into another bar because of small number of observations. Therefore, three ranges can be considered for Fig. 4(a); lower range (i.e., 0–3000 vehicles/hour), mid-range (i.e., 3000–4000 vehicles/hour), and upper range (i.e., 4000–6000 vehicles/hour). As observed from Fig. 4(a), first of all, no consistent positive or negative trend is found between the percentages of "Crash code 1" and varying ranges of traffic counts during the 1-hour after period of freeway crash occurrences; the maximum percentage of "Crash code 1" is found for the mid-range (i.e., 3000–4000 vehicles/hour). Second, lower percentages of "Crash code 1" during the after period of ASCS implementation compared to the before period of ASCS implementation are found for middle (i.e., 3000–4000 vehicles/hour) and upper ranges (i.e., 4000–6000 vehicles/hour). Therefore, it can be concluded that the favorable effect of ASCS found for Charleston

Table 9
Fixed-parameter logistic regression model estimates & interpretations of after-period indicator of ASCS deployment.

Corridor type	Corridor Name	Coefficients:		Odds ratio	% change in the secondary crash occurrence odds
		Estimate	Pr(>|z|)		
Freeways with ASCS deployed on alternate routes	Charleston I-26 E	−0.641	0.059*	0.527	−47.324%
	Charleston I-26 W	0.222	0.5182	-	-
Freeways with non-ASCS alternate routes	Richland-Lexington I-26 E	−0.065	0.8518	-	-
	Richland-Lexington I-26 W	0.334	0.3030	-	-

'*' statistically significant at a 0.1 significance level.

Table 10
Fixed-parameter logistic regression model estimates of other predictors.

Predictors	Coefficients estimate (with Pr(>|z|) in the parentheses)			
	Freeways with ASCS deployed on alternate routes		Freeways with non-ASCS alternate routes	
	Charleston I-26 E	Charleston I-26 W	Richland-Lexington I-26 E	Richland-Lexington I-26 W
Light condition	NS	0.378 (0.093*)	0.599 (0.028**)	NS
Weather condition	NS	0.547 (0.015**)	NS	NS
Rear end	0.821 (0.0008**)	1.417 (4.96e-07**)	0.949 (0.0003**)	1.077 (3.72e-05**)
Angle crash	NS	1.397 (0.0006*)	NS	NS
Weekday	NS	NS	0.498 (0.068*)	−0.912 (4.4e-05**)
Peak period	0.929 (7.32e-06**)	NS	NS	NS
Temporal trend	0.268 (0.041**)	NS	NS	NS
log(AADT)	1.916 (0.0001**)	NS	1.876 (0.001**)	NS

'**' statistically significant at a 0.05 significance level.
'*' statistically significant at a 0.1 significance level.
'NS' not statistically significant.

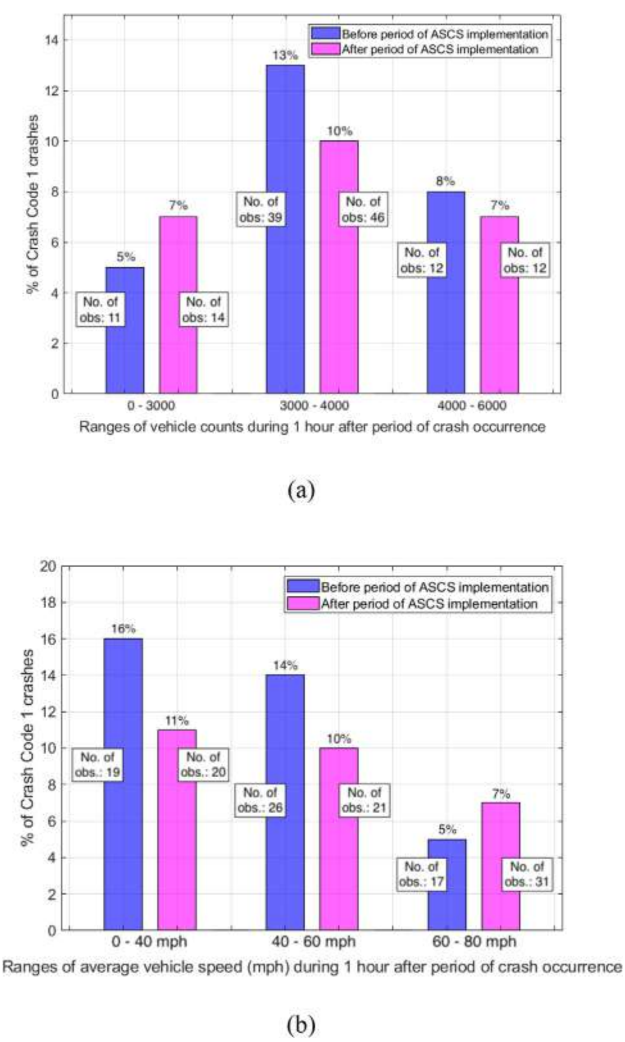


Fig. 4. Bar charts showing percentages of Crash code 1 across ranges of (a) traffic counts and (b) average speed on the freeway during the 1-hour after period of freeway crash occurrences.

I-26 E is not contributed by reduced exposure or low traffic counts on the freeway (i.e., traffic counts on the freeway can be lower than usual as the freeway drivers start to take the alternate route after a crash occurrence on the freeway).

However, in Fig. 4(b), it is observed that there is a negative or downward trend between the percentages of "Crash code 1" and the ranges of average vehicle speed during the 1-hour after period of freeway crash occurrences. While it is pretty much intuitive for primary crashes that increased speed on the freeway could cause higher percentage of primary crashes (Abdel-Aty et al., 2007), same cannot be stated with confidence for the freeway secondary crashes. It is to be noted that, the freeway secondary crashes are caused due to sudden congestion/queue on the freeway because of sudden lane blockage or disturbance in the traffic flow caused by a primary crash on the freeway. Thus, higher average speed on the freeway during the 1-hour after period of a freeway crash occurrence indicates less impact due to that crash. On the contrary, lower average speed on the freeway during the 1-hour after period of a freeway crash occurrence indicates higher disturbance caused by that crash, which would increase the risk for a secondary crash.

To validate this further, the authors also perform a separate logistic regression modeling (as shown in Table 11) utilizing the traffic counts and average speed information during the 1-hour after period of freeway crash occurrences as continuous variables. As the "traffic counts" variable is not found significant at a 0.1 significance level, the authors do not include it and other insignificant variables from the model presented in Table 11 in order to obtain a better fit. It should be mentioned that, these two variables were not used in the models presented earlier in this paper because the traffic counts and average speed information is not available for all the crashes. The authors found that, the traffic counts and average speed information is available for about 95% of the crash data considered for Charleston I-26 E. However, just for the sake of investigating if there is any relationship between the likelihood of secondary crashes and traffic counts or average speed during the 1-hour after period of freeway crash occurrences, the available data can be considered sufficient. As shown in Table 11, the "Speed" variable (i.e., average speed on the freeway during the 1-hour after period of a freeway crash occurrence) is found to be significant at a 0.05 significance level and it has a negative coefficient which further validates that as the average speed on the freeway is

Table 11
Estimates of the random-parameter logistic regression model using speed as an explanatory variable (for Charleston I-26 E).

Predictors	Coefficients	
	Estimate	Pr(>|z|)
(Intercept)	−507.5	0.061*
Rear-end	0.911	5.02e-05**
After-period indicator of ASCS deployment	−0.588	0.089*
Temporal trend	0.239	0.075*
Speed	−0.021	2.80e-04**
log(AADT)	1.810	6.56e-04**

'*' statistically significant at a 0.1 significance level.
'**' statistically significant at a 0.05 significance level.

higher during the 1-hour after period of a freeway crash occurrence, the likelihood of secondary crash occurrence is lower (same as concluded based on Fig. 4(b)). Therefore, the authors conclude that the favorable effect of ASCS found for Charleston I-26 E in terms of reducing the likelihood of secondary crashes is not a contribution of reduced crash exposure (i.e., lower freeway traffic counts after a crash occurrence) or reduced speed on the freeway.

5.3. Random-parameter logistic regression model results

Fixed-parameter binary logistic regression models help to identify the statistically significant variables in the likelihood of secondary crash. However, a limitation of fixed-parameter binary logistic regression model is that it cannot capture unobserved heterogeneity since the model assumes the global effect of each predictor across observations. Therefore, a random parameter logistic regression model is deployed to study the heterogeneous effect of the ASCS deployment for Charleston I-26 E. The results for Charleston I-26 E is of interest and thus presented here since it shows statistically significant ASCS effect in the previous section using fixed-parameter binary regression modeling.

The estimations of random-parameter logistic model for Charleston I-26 E is presented in Table 12. The random-parameter model explains the variability in the effect of ASCS deployment (i.e., ASCS deployment variable) across observations and provides more significant parameters over the fixed-parameter model (e.g., angle crash variable becomes significant in the random-parameter model). Although the weekday variable is not significant in the random-parameter model, it is still kept since keeping the weekday variable reduces the AIC value and improves the overall goodness of fit of the model as observed in our analyses. Note that the fixed-parameter logistic regression model presented in Table 12 has 2 additional variables (i.e., angle crash, and weekday) compared to the model presented previously in Table 10. The additional variables are included to compare the goodness of fits between the fixed-parameter logistic regression model and the random-parameter logistic regression model. The likelihood ratio test suggests that the random-parameter model improves the overall goodness of fit of the model compared to the fixed-parameter model, as shown in Table 13. The standard deviation associated with the presence of ASCS (i.e., S.D. ASCS) is statistically significant at a 0.05 significance level, indicating the presence of unobserved heterogeneity across observations.

As indicated in Table 12, the random parameter of ASCS follows a normal distribution with a mean of −2.305 and a standard deviation of 2.351. Since the parameter of ASCS follows the normal distribution, it is estimated that 84% of all observations have a negative coefficient associated with the presence of ASCS corridor, suggesting an association between the presence of the ASCS deployed on the alternate route and the reduction of the likelihood of secondary crashes on the parallel freeway. For the remaining 16% of all observations, the coefficients associated with the presence of ASCS deployed on the alternate route are positive, suggesting an association between the presence of the ASCS deployed on the alternate route and the increase of the likelihood of freeway secondary crashes.

Fig. 5 shows the kernel density of the individual's conditional means for the coefficient of ASCS. It turns out that the majority of the individual's conditional means (the unshaded portion in the Figure) has negative signs, suggesting the presence of ASCS associated with reductions of the likelihood of freeway secondary crashes for most of the observations.

The authors then investigate the locations of individual observations (i.e., primary crashes) for which the presence of the ASCS deployed on alternate route is associated with an increase in the likelihood of freeway secondary crashes. Fig. 6 shows the locations of the crashes on Charleston I-26 E (in blue color on the freeway) and the possible exit ramps to exit Charleston I-26 E to access US 52. As observed from Fig. 6, most of the primary crashes for which coefficients of ASCS are positive occurred closer to the east end of the Charleston I-26 E section and took place past the second possible exit ramp to access US 52. Also, closer to the east end of the Charleston I-26 E section means closer to the Charleston city downtown. When a crash occurs closer to the east end of Charleston I-26 E section, it may not always seem to be a convenient choice for the upstream traffic to divert as they

Table 12
Results of model estimation for Charleston I-26 E with ASCS deployed on alternate route.

Predictors	Coefficients			
	Fixed-parameter logistic regression model		Random-parameter logistic regression model	
	Estimate	Pr (>|z|)	Estimate	Pr (>|z|)
Constant	−26.209	1.73E-05 **	−31.396	6.73e-05 **
Temporal trend	0.265	0.0439 **	0.3230	0.0433 **
Rear end	0.802	0.001 **	1.087	0.002 **
Angle crash	0.679	0.117	0.959	0.082*
Weekday	−0.370	0.302	−0.628	0.168
Peak period	0.916	8.63E-06 **	1.107	3.06E-05 **
log(AADT)	1.916	0.0002 **	2.334	0.0003**
Mean. ASCS	−0.631	0.063*	−2.305	0.06*
S.D. ASCS	NA	NA	2.351	0.0247**

'**' statistically significant at a 0.05 significance level.
'*' statistically significant at a 0.1 significance level.
'NA' not available for the fixed-parameter model.

Table 13
Likelihood ratio tests results.

	Degrees of freedom	Log-Likelihood	Difference in degrees of freedom	Chisq	Pr(>Chisq)
Fixed-parameter logistic regression model	8	−450.22			
Random-parameter logistic regression model	9	−448.53	1	3.3736	0.066*

'*' statistically significant at a 0.1 significance level.

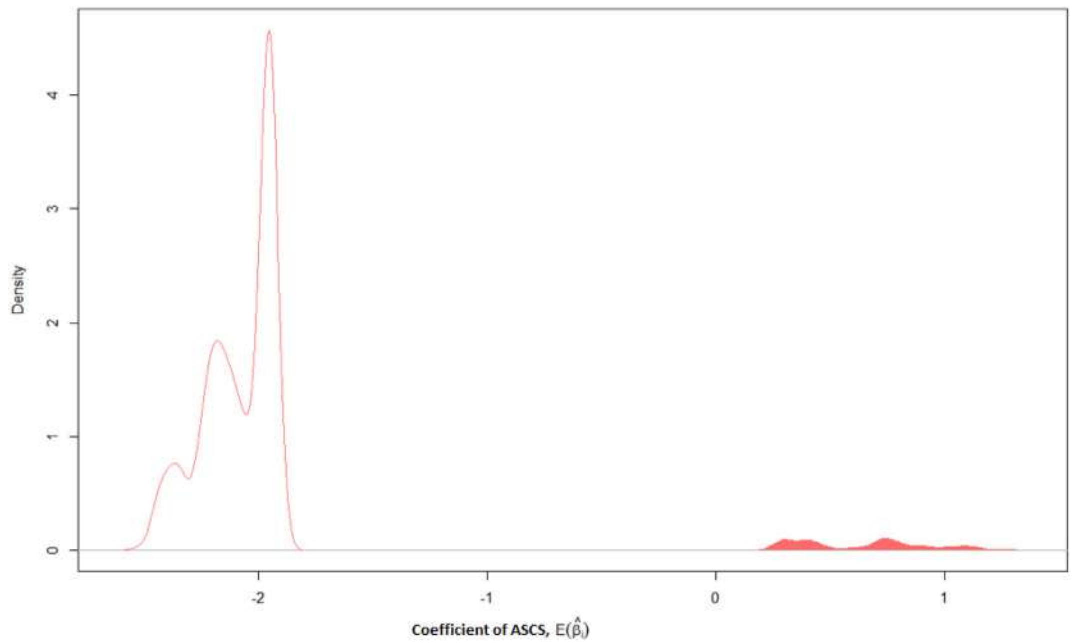


Fig. 5. Kernel density of the individual's conditional means for the coefficient of ASCS (Charleston I-26 E with ASCS deployed on alternate route US 52).

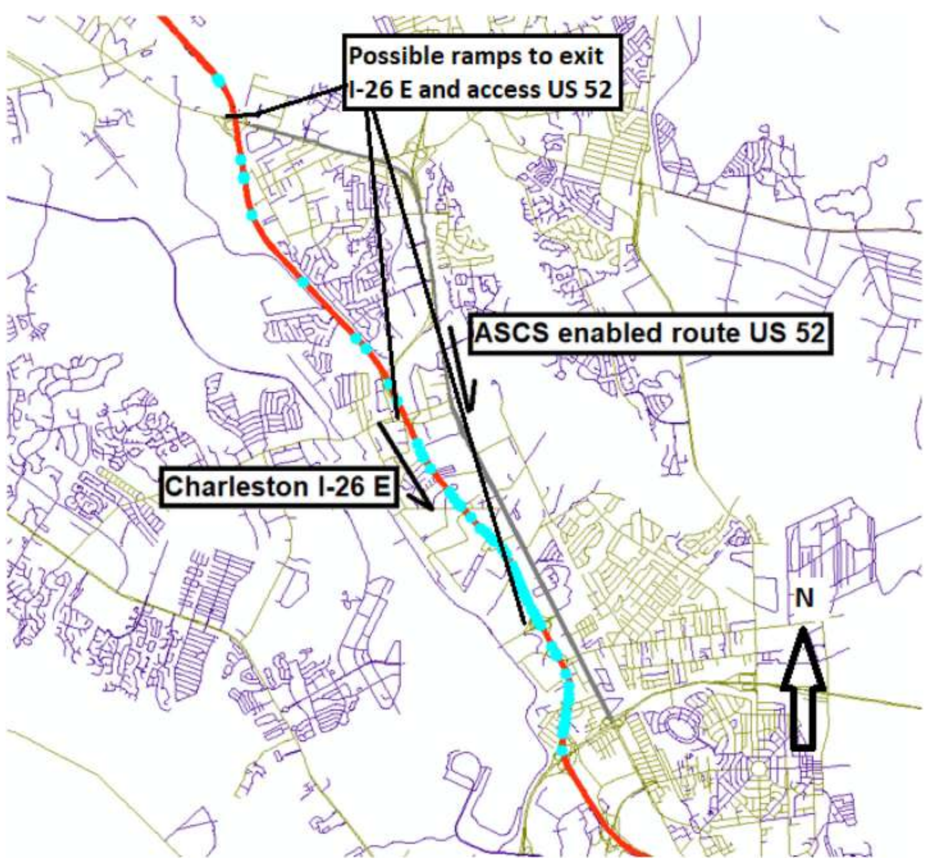


Fig. 6. Location of Charleston I-26 E freeway crashes associated with the increase in the likelihood of secondary crashes.

may think that they are already very close to their destination and sometimes it may not be a feasible option to divert as they may have already passed the nearest exit ramp. Also, the benefit of ASCS deployed on the alternate route may not be that much since fewer signals are involved when a crash occurs closer to the east end of Charleston I-26 E section. Therefore, the effect of ASCS deployed on alternate route on the likelihood of freeway secondary crashes can vary depending on the location of the primary freeway crash as it affects the amount of traffic are that is able or chooses to divert.

6. Conclusions

The reduction of the likelihood of secondary crashes can noticeably decrease the emissions, delays, vehicle operating costs, and safety issues on the freeways. This research unveils a unique interrelation between ASCS deployed on alternate route and the likelihood of parallel freeway secondary crashes. The findings from a fixed-parameter binary logistic regression model using 52 months crash data of Charleston I-26 E with ASCS deployed on alternate route US 52 shows a 47% reduction of the likelihood of freeway secondary crashes. The authors further investigate Charleston I-26 E using a random-parameter binary logistic regression model to find unobserved heterogeneity and find that 84% of all observations have negative coefficients associated with the presence of ASCS deployed on alternate route, suggesting an association between the presence of ASCS deployed on alternate route and the reduction of the likelihood of secondary crashes on the parallel freeway. The benefit of ASCS deployment on an alternate route towards freeway secondary crash reduction is found to be dependent on the location of the primary crash as the location of the primary crash determines how much of the upstream traffic will be able or choose to take the exit ramp to the ASCS deployed alternate route. Therefore, the findings provide a new insight on improving safety on a freeway with the implementation of ASCS arterials that could be used as alternate routes during an incident on the freeway. The results also reveal that other contributing factors such as rear-end crash, light condition, peak period, weekday and AADT increase the likelihood of secondary crashes on a freeway.

In this research, the authors have identified secondary crashes on freeways based on a predefined spatio-temporal criterion adopted extensively in previous work. Selecting a fixed spatio-temporal criterion for identifying the secondary crashes on the freeways is a challenging task. For some primary crashes that occur on the freeways, the corresponding secondary crashes can sometimes occur beyond a predefined spatio-temporal range since the impacts of some primary crashes on traffic can exceed the predefined spatio-temporal criterion. In the future, further research will be conducted considering various spatio-temporal criteria. In addition, due to limitation of appropriate data on the drivers' safety awareness (South Carolina Traffic Collision Report Form (TR-310), 2012), the authors did not consider the safety awareness of drivers as a factor while modeling the likelihood of secondary crashes. Unavailability of lane blockage and clearance time information for individual crash is another limitation of this study. Also, different ASCS controllers follow different types of algorithms. For this study, all the intersections on the alternate routes considered are equipped with only one type of ASCS technology. However, our future research will focus on identifying how different types of ASCS deployed on the alternate routes can affect the freeway secondary crash occurrences.

7. Practical applications

Analysis results indicate that, there is an association between ASCS deployed on an alternate route and the likelihood of secondary crashes on the parallel freeway. Therefore, it is recommended that the SCDOT considers utilizing ASCS on corridors that are often used as alternate routes when there is a crash on the adjacent parallel freeways. According to the findings of this paper, existence of such an ASCS-deployed alternate route can help reduce the likelihood of freeway secondary crashes and can improve freeway safety.

Declarations of competing interest

None.

Acknowledgments

The authors acknowledge the South Carolina Department of Transportation (SCDOT), which provided funding (Research Project No.: SPR 737) for this research. The authors also sincerely acknowledge SCDOT SPR 737 Steering Committee for providing valuable guidance to the authors for this research. The contents of this paper reflect the views of the authors who are responsible for the facts and the accuracy of the presented data. The contents do not reflect the official views of SCDOT or FHWA. This paper does not constitute a standard, specification, or regulation.

References

Abdel-Aty, M., Pande, A., Lee, C., Gayah, V., & Santos, C. D. (2007). Crash risk assessment using intelligent transportation systems data and real-time intervention strategies to improve safety on freeways. *Journal of Intelligent Transportation Systems, 11*(3), 107–120. https://doi.org/10.1080/15472450701410395.

Chowdhury, M. A., Fries, R., Ma, Y., Dunning, A. E., Hamlin, C., Ozbay, K., . . ., Sarasua, W. (2007). Benefit cost analysis of accelerated incident clearance (FHWA-SC-07-01). Article FHWA-SC-07-01. https://trid.trb.org/view/850224.

Eghtedari, A. G. (2006). Measuring benefits of adaptive traffic signal control: Case study of mill plain Boulevard, Vancouver, Washington (No. 06–0111). Article 06–0111. Transportation Research Board 85th Annual Meeting. Transportation Research Board. https://trid.trb.org/view/776184.

Elkins, S., & Niehus, G. (2012). InSync adaptive traffic control system for the veterans memorial Hwy Corridor on Long Island, NY (dot:24966). C-10–01. https://rosap.ntl.bts.gov/view/dot/24966.

Fontaine, M. D., Ma, J., & Hu, J. (2015). Evaluation of the Virginia Department of Transportation adaptive signal control technology pilot project. (dot:29013). VCTIR 15-R24. https://rosap.ntl.bts.gov/view/dot/29013.

Gartner, N. H., Pooran, F. J., & Andrews, C. M. (2002). Optimized policies for adaptive control strategy in real-time traffic adaptive control systems: Implementation and field testing. *Transportation Research Record, 1811*(1), 148–156. https://doi.org/10.3141/1811-18.

Goodall, N. J. (2017). Probability of secondary crash occurrence on freeways with the use of private-sector speed data. *Transportation Research Record, 2635*(1), 11–18. https://doi.org/10.3141/2635-02.

Hirunyanitiwattana, W., & Mattingly, S. P. (2006). Identifying secondary crash characteristics for California Highway System (No. 06–2900). Article 06–2900. Transportation Research Board 85th Annual Meeting. Transportation Research Board. https://trid.trb.org/view/777729.

JJin, W., Khan, S. M., & Chowdhury, M. (2019). Exploring the Safety Effects of Adaptive Signal Control Systems. International Conference on Transportation and Development 2019: Smarter and Safer Mobility and Cities, 8–18. American Society of Civil Engineers. https://doi.org/10.1061/9780784482575.002.

Jin, W., Chowdhury, M., Khan, S. M., & Gerard, P. (2020). Investigating the impacts of crash prediction models on quantifying safety effectiveness of Adaptive Signal Control Systems. Journal of Safety Research. https://doi.org/10.1016/j.jsr.2020.11.003.

Jin, W., Chowdhury, M., Salek, M. S., Khan, S. M., & Gerard, P. (2021). Investigating hierarchical effects of adaptive signal control system on crash severity using random-parameter ordered regression models incorporating observed heterogeneity. Accident Analysis & Prevention, 150, 105895. https://doi.org/10.1016/j.aap.2020.105895.

Karlaftis, M. G., Latoski, S. P., Richards, N. J., & Sinha, K. C. (1999). ITS impacts on safety and traffic management: An investigation of secondary crash causes. *ITS Journal - Intelligent Transportation Systems Journal, 5*(1), 39–52. https://doi.org/10.1080/10248079908903756.

Kergaye, C., Stevanovic, A., & Martin, P. T. (2009). Comparison of before-after versus off–on adaptive traffic control evaluations: Park City, Utah,case study. *Transportation Research Record, 2128*(1), 192–201. https://doi.org/10.3141/2128-20.

Khattak, Z. H. (2016). *Evaluating the operational and safety aspects of adaptive traffic control systems in Pennsylvania [University of Pittsburgh ETD]*. University of Pittsburgh.

Kitali, A. E., Alluri, P., Sando, T., Haule, H., Kidando, E., & Lentz, R. (2018). Likelihood estimation of secondary crashes using Bayesian complementary log-log model. *Accident Analysis & Prevention, 119*, 58–67. https://doi.org/10.1016/j.aap.2018.07.003.

Kitali, A. E., Alluri, P., Sando, T., & Wu, W. (2019). Identification of secondary crash risk factors using penalized logistic regression model. *Transportation Research Record, 2673*(11), 901–914. https://doi.org/10.1177/0361198119849053.

Latoski, S. P., Pal, R., & Sinha, K. C. (1999). Cost-effectiveness evaluation of hoosier helper freeway service patrol. *Journal of Transportation Engineering, 125*(5), 429–438. https://doi.org/10.1061/(ASCE)0733-947X(1999)125:5(429).

Lowrie, P. R. (1990). SCATS, Sydney co-ordinated adaptive traffic system: A traffic responsive method of controlling urban traffic. https://trid.trb.org/view/488852.

Mirchandani, P., & Head, L. (2001). A real-time traffic signal control system: Architecture, algorithms, and analysis. *Transportation Research Part C: Emerging Technologies, 9*(6), 415–432. https://doi.org/10.1016/S0968-090X(00)00047-4.

Moore, J. E., Giuliano, G., & Cho, S. (2004). Secondary accident rates on Los Angeles Freeways. *Journal of Transportation Engineering, 130*(3), 280–285. https://doi.org/10.1061/(ASCE)0733-947X(2004)130:3(280).

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity, 41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6.

Persaud, B., Lan, B., Lyon, C., & Bhim, R. (2010). Comparison of empirical Bayes and full Bayes approaches for before–after road safety evaluations. *Accident Analysis & Prevention, 42*(1), 38–43. https://doi.org/10.1016/j.aap.2009.06.028.

Raub, R. A. (1997a). Secondary crashes: an important component of roadway incident management. *Transportation Quarterly, 51*(3) https://trid.trb.org/view/577266.

Raub, R. A. (1997b). Occurrence of secondary crashes on Urban Arterial Roadways. *Transportation Research Record, 1581*(1), 53–58. https://doi.org/10.3141/1581-07.

Sarrias, M. (2016). Discrete choice models with random parameters in *R*: The rchoice package. *Journal of Statistical Software, 74*(10). https://doi.org/10.18637/jss.v074.i10.

SCDOT Iteris ClearGuide. Available at https://scdot.iteris-clearguide.com/. Accessed on September 14, 2020.

SCDOT SHEP. Available at: https://www.scdot.org/travel/travel-emergencyresponse.aspx Accessed on September 14, 2020.

SCDOT Traffic Counts. Available at https://www.scdot.org/travel/travel-trafficdata.aspx. Accessed on September 14, 2020.

So, J., Stevanovic, A., Posadas, E., & Awwad, R. (2014). Field evaluation of a SynchroGreen adaptive signal system. *Second Transportation & Development Congress, 2014*, 388–399. https://doi.org/10.1061/9780784413586.037.

South Carolina Traffic Collision Report Form (TR-310). Available at: https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/sc_tr310_manual_rev8_2012.pdf#:~:text=Each%20TR%2D310%20consists%20of,of%20Highway%20Safety%20(OHS). Accessed on September 14, 2020.

Washington, S., Karlaftis, M. G., Mannering, F., & Panagiotis, A. (2020). *Statistical and econometric methods for transportation data analysis*. CRC Press.

Xu, C., Liu, P., Yang, B., & Wang, W. (2016). Real-time estimation of secondary crash likelihood on freeways using high-resolution loop detector data. *Transportation Research Part C: Emerging Technologies, 71*, 406–418. https://doi.org/10.1016/j.trc.2016.08.015.

Yang, H., Bartin, B., & Ozbay, K. (2013). Use of sensor data to identify secondary crashes on freeways. *Transportation Research Record*. https://doi.org/10.3141/2396-10.

Yang, H., Ozbay, K., & Xie, K. (2014). Assessing the risk of secondary crashes on highways. *Journal of Safety Research, 49*, 143.e1. https://doi.org/10.1016/j.jsr.2014.03.007.

M Sabbir Salek is a Ph.D. student at Clemson University, South Carolina. He received his B.Sc. in Mechanical Engineering from Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh. His research interest includes data analytics, regression modeling, machine learning, connected and automated vehicles and trucks, and cloud computing.

Weimin Jin is a Ph.D. candidate at Clemson University. He received his M.Sc. in 2015 from Tongji University in China. His research interest includes data analytics, regression modeling and inference, probability, Bayesian modeling and inference, traffic safety, connected and automated vehicles, intelligent transportation systems, and traffic simulation.

Dr. Sakib Mahmud Khan is the assistant director of the Center for Connected Multimodal Mobility (C_2M_2). Before joining the center he was a post-doctoral research scholar working at California Partners for Advanced Transportation Technology (PATH), University of California Berkeley. He received his Ph.D. and M. Sc. in Civil Engineering from Clemson University in 2019 and 2015, respectively.

Dr. Mashrur Chowdhury is the Eugene Douglas Mays Professor in the Glenn Department of Civil Engineering at Clemson University. He is the Director of the USDOT Center for Connected Multimodal Mobility (C2M2) (a TIER 1 USDOT University Transportation Center). He is Co-Director of the Complex Systems, Data Analytics and Visualization Institute at Clemson University. He is a Fellow of the American Society of Civil Engineers. He is a registered professional engineer in Ohio.

Dr. Patrick Gerard is a Professor in the School of Mathematical and Statistical Sciences at Clemson University. He currently teaches courses in Multivariate Statistics, Environmental Statistics, and Spatial Statistics. His research includes collaborative efforts with faculty and staff in a wide range of disciplines as well as statistical methodological research primarily in the areas of nonparametric function estimation and computational methods in statistics.

Saurabh Bikram Basnet is a M.S student at Clemson University. His research interests include traffic safety and crash data analysis using statistical models. He also has interests in traffic operations, simulation and intelligent transportation systems. He currently teaches Geomatics to sophomore students at Clemson University.

Mohammad Torkjazi is currently working as Senior Operations Research Analyst at FedEx Express. He is also a PhD candidate of Transportation Engineering at the University of South Carolina. Mohammad received his M.S. from Sharif University of Technology, Iran and his B.S. from Azad University, Iran. His current research is to investigate the congestion surrounding container terminals (air, marine, rail) by leveraging a blend of optimization, machine learning, simulation, and programming skills. Mohammad designed optimization models to facilitate the truck appointment reservation at marine container terminals.

Dr. Nathan Huynh is a Professor in the Department of Civil and Environmental Engineering at the University of South Carolina (USC) with specialization in transportation engineering. Prior to coming to USC in 2008, he worked at the Port of Houston Authority as a project manager and systems analyst. He received both of his Ph.D. and M.S. degrees in Transportation Engineering from the University of Texas at Austin. His current research interests include 1) civil infrastructure systems, 2) freight transportation systems, 3) transportation and public health, and 4) healthcare systems reengineering.

Journal of Safety Research 76 (2021) 205–217

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Assessment of temporal stability in risk factors of crashes at horizontal curves on rural two-lane undivided highways

Nabeel Saleem Saad Al-Bdairi [a,*], Ali Behnood [b]

[a] College of Engineering, Civil Engineering Department University of Wasit, Kut, Iraq
[b] Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall, West Lafayette, IN 47907-2051, USA

ARTICLE INFO

Article history:
Received 14 April 2020
Received in revised form 16 July 2020
Accepted 3 December 2020
Available online 31 December 2020

Keywords:
Injury severity
Two-lane undivided highways
Temporal stability
Horizontal curve
Heterogeneity in means and variances

ABSTRACT

Introduction: Safety of horizontal curves on rural two-lane, two-way undivided roadways is not fully explored. This study investigates factors that impact injury severity of such crashes. *Method:* To achieve the aim of this paper, issues associated with police-reported crash data such as unobserved heterogeneity and temporal stability need to be accounted for. Hence, a mixed logit model was estimated, while heterogeneity in means and variances is investigated by considering four injury severity outcomes for drivers: severe injury, moderate injury, possible injury, and no injury. Crash data for the period between 2011 and 2016 for crashes that occurred in the state of Oregon was analyzed. Temporal stability in factors determining the injury severity was investigated by identifying three time periods through splitting crash data into 2011–2012, 2013–2014, and 2015–2016. *Results:* Despite some factors affecting injuries in all specified time periods, the values of the marginal effects showed relative differences. The estimation results revealed that some factors increased the risk of being involved in severe injury crashes, including head-on collisions, drunk drivers, failure to negotiate curves, older drivers, and exceeding the speed limits. *Conclusions:* The hypothesis that attributes of injury severity are temporally stable is rejected. For example, young drivers (30 years old and younger) and middle-aged drivers were found to be temporally instable over time. *Practical applications:* The findings could help transportation authorities and safety professionals to enhance the safety of horizontal curves through appropriate and effective countermeasures.

1. Introduction

Horizontal curves are integral elements in roadway infrastructure since they play a vital role in changing the alignment or the direction of roadways. Despite the significant role of these curves in roadway design, such locations are still safety hazards to road users in the United States. Therefore, enhancing safety at horizontal curves has been set as an overarching goal for transportation agencies across the United States. However, this is not an easy task because some factors are commonly associated with horizontal curves. These factors include increased demand on drivers (Calvi, 2015; Pratt et al., 2018), underestimation of curve sharpness (Pratt et al., 2018; Schneider et al., 2009), curve length (Rakotonirainy et al., 2015), curve radii (Geedipally et al., 2019), sight distance (Rakotonirainy et al., 2015), speed limit (Wu et al., 2019), poor signage (Rakotonirainy et al., 2015), inclement weather and its impact on side friction (Pratt et al., 2018; Rakotonirainy et al., 2015), and superelevation (Geedipally et al., 2019).

It has been reported that the number of fatalities occurring due to crashes along horizontal curves is higher compared to that on tangent segments, despite the disproportionate length of horizontal curves in the roadway network (Geedipally et al., 2019; Lord et al., 2011; Torbic et al., 2004; Wu and Xu, 2017). Nationwide, recent statistics show that the second most common vehicle maneuver before a fatal crash was negotiating a curve, with nearly 18.5% of all reported fatal crashes in 2017 (NHTSA, 2019). The same trend has also been seen statewide in Oregon, in which 27.3% (110 out of 403) of the reported fatalities in 2017 occurred on horizontal curves (ODOT, 2019). It is worth noting that crashes along rural two-lane undivided highway (RTU) horizontal curves encompass the majority of reported crashes at horizontal curves.

In light of this, highway engineers, transportation agencies, and safety researchers strive to leverage safety at RTU horizontal curves to improve the overall safety. This could be accomplished by capturing determinants of injury severity that resulted from crashes along RTU horizontal curves. By doing so, appropriate countermeasures could be implemented. However, this depends

* Corresponding author.
E-mail addresses: nsaleem@uowasit.edu.iq (N.S.S. Al-Bdairi), abehnood@purdue.edu (A. Behnood).

https://doi.org/10.1016/j.jsr.2020.12.003
0022-4375/

on the performance and accuracy of developed methodological approaches in terms of their inferences to make accurate predictions because erroneous and inefficient interventions can be implemented if inaccurate methods are used. This study attempts to address notorious limitations in traffic safety analysis methods, namely, unobserved heterogeneity and temporal stability of contributing factors.

A major gap in the literature regarding the safety of horizontal curves is ignoring temporal stability and unobserved heterogeneity. As such, the current study aims to fill this gap. By doing so, a better understanding of determinants of injury severity from these crashes can be gained. Achieving this overarching goal could assist transportation agencies and safety professionals to enhance safety at horizontal curves through appropriate countermeasures targeting, in part, geometric design features of the curves and/or emphasizing law enforcement.

2. Literature review

Compared with highway tangents, RTU horizontal curves pose serious concerns to traffic safety. As such, an abundance of research has been devoted to rigorously emphasize horizontal curve safety from different perspectives, including safety–critical events on rural two-lane curves (Wang et al., 2017), crash frequency (Dhahir & Hassan, 2019a, 2019b; Gooch et al., 2016; Xin et al., 2017), visibility of horizontal curve (Jamson et al., 2015), developing safety performance functions (Banihashemi, 2016; Gooch et al., 2018), reliability analysis (Jesna & Anjaneyulu, 2016), run-off risk (Choudhari & Maji, 2019), motorcycle safety (Gabauer & Li, 2015; Xin et al., 2019), large truck safety (Fitzsimmons et al., 2012), injury severity (Schneider et al., 2009), crash frequency by severity (Anarkooli et al., 2019), and drivers' speed behavior when negotiating horizontal ramp curves in interchanges (Farah et al., 2019).

This literature shows that exploring injury severity of crashes along RTU horizontal curves is highly overlooked, with only research conducted by Schneider et al. (2009). However, this research failed to address unobserved heterogeneity issue. This is because the multinomial logit model used in their study was a very promising model at that time when there was no such advancement in methodological methods. Therefore, the current study aims at overcoming the limitations highlighted in the literature by using advanced methods that are capable of capturing unobserved heterogeneity, which is more likely to exist in police-reported crash reports (Mannering et al., 2016). Recently, examining unobserved heterogeneity used the random parameters approach, without exploring whether the means and variances of potential random parameters being a function of explanatory variable or not is criticized (Al-Bdairi et al., 2020; Alnawmasi & Mannering, 2019; Behnood & Mannering, 2017a; Seraneeprakarn et al., 2017). In addition to this issue, a temporal stability concern also arises in the crash data, which was highly ignored in past research that emphasized crashes along horizontal curves (Anarkooli et al., 2019; Choudhari & Maji, 2019; Mannering, 2018; Schneider et al., 2009; Wang et al., 2017). Taken together, the current study investigated temporal stability of factors impacting injury severity of drivers involved in crashes along RTU horizontal curves via estimating mixed logit models with heterogeneity in means and variances approach and splitting the crash data into three time periods.

3. Methodological approach

The so-called heterogeneity models have been widely used in the analysis of crash-injury severity due to their superior performance compared to other models (Al-Bdairi, 2020; Al-Bdairi et al., 2018; Al-Bdairi & Hernandez, 2017; Behnood et al., 2014; Behnood & Mannering, 2017b, 2017c). Despite their superiority in model fitting, unobserved heterogeneity models have also some drawbacks (Mannering et al., 2016). For instance, the random parameters approach fails to investigate the relationship between the unobserved heterogeneity and explanatory variables. In other words, the means and variances of the random parameters are fixed (homogeneous) across the observations. However, this is not the case if an analyst seeks more flexibility in the developed models. That being said, the random parameters approach with heterogeneity in means and variances is utilized in the current study to overcome the aforementioned shortcomings in the most commonly used methods in recent years, and also to come up with efficient, unbiased, and more accurate inferences.

Following Milton et al. (2008) and Washington et al. (2011), an injury severity function T_{kj}should be firstly introduced, and this function is written as:

$$T_{kj} = \boldsymbol{\beta_k X_{kj}} + \varepsilon_{kj} \tag{1}$$

where T_{kj} is a function of injury severity determining the driver severity k in crash j, $\boldsymbol{\beta}_k$ is the vector of the estimated coefficients for injury k, $\boldsymbol{X}_{kj}$ is the associated vector of attributes affecting driver injury severity k resulted from crash j, and ε_{kj} is the error term. It should be noted that the error term follows a generalized extreme value distribution. Accounting for heterogeneity in the means and variances of random parameters is illustrated in Eq. (2). This should be achieved through letting $\boldsymbol{\beta}_{kj}$ as a vector of estimated parameters, in which parameters vary across crashes (Alnawmasi & Mannering, 2019; Behnood & Mannering, 2019; Behnood and Mannering, 2017a; Seraneeprakarn et al., 2017).

$$\boldsymbol{\beta}_{kj} = \beta_k + \Theta_{kj}\boldsymbol{Z}_{kj} + \sigma_{kj} EXP(\boldsymbol{\omega}_{kj}\boldsymbol{W}_{kj})\,\upsilon_{kj} \tag{2}$$

where β_k is the mean parameter, $\boldsymbol{Z}_{kj}$ is a vector of attributes that accommodate the heterogeneity in the mean,Θ_{kj} is a corresponding vector of estimable parameters, $\boldsymbol{W}_{kj}$is a vector of attributes that addresses heterogeneity in the standard deviation σ_{kj}, $\boldsymbol{\omega}_{kj}$is the corresponding parameter vector, and $\upsilon_{kj}$is a randomly distributed term that accounts for unobserved heterogeneity across RTU horizontal crashes. To derive the probability of a crash j results in driver injury severity k, the framework presented by McFadden and Train (2000) and Washington et al. (2011) should be followed as:

$$P_j(k) = \int \frac{EXP(\boldsymbol{\beta_k X_{kj}})}{\sum_{\forall K} EXP(\boldsymbol{\beta_k X_{kj}})} f(\boldsymbol{\beta}|\boldsymbol{\varphi})d\boldsymbol{\beta} \tag{3}$$

where $f(\boldsymbol{\beta}|\boldsymbol{\varphi})$ is the density function of $\boldsymbol{\beta}$ with $\boldsymbol{\varphi}$ denoting vector of parameters (mean and variance) of density function. A simulated maximum likelihood with 200 Halton draws has been utilized herein in estimating the models (McFadden & Train, 2000). Also, extensive parametric distributional forms, including normal, lognormal, triangular, and uniform have been tried for random parameters. Yet, only normal distribution yields a superior statistical fit. The marginal effects are also calculated to ease interpretation of the estimated results. The difference in the estimated probabilities given that indicator variables are altered from zero to one represents the marginal effects (Washington et al., 2011).

4. Empirical setting

In this study, crash data were drawn from Oregon Department of Transportation (ODOT) involving crashes that occurred along RTU horizontal curves in Oregon for the period between 2011 and 2016. After removing incomplete or crashes with missing information, a total of 13,882 observations are obtained in the final dataset. The final dataset includes a great deal of crash information on driver attributes (i.e., driver sobriety, gender, age, seatbelt usage, driving license status, driver fatigue, and if driver is distracted); roadway characteristics (i.e., roadway surface conditions, traffic control device, and speed limit); crash attributes (i.e., collision type, number of vehicles involved, and major cause of crash); vehicle features (type of involved vehicle); time-related attributes (such as time of crash); and weather and lighting conditions.

To overcome a limitation in the final dataset represented by lower observations of fatal crashes, injury severity of drivers was collapsed into four main groups: severe injury (fatal and incapacitating injury), moderate injury (non-incapacitating), possible injury (possible injury), and no injury (property damage only). Since this study sought to statistically assess the temporal instability in the risk factors of crashes along RTU horizontal curves, final dataset was split into three time periods[1]: 2011–2012, 2013–2014, and 2015–2016. The distribution of injury severity levels in the final data set in the four time periods is provided in Table 1. Further, the final data set includes numerous variables that affect driver injury severity resulted from RTU horizontal curves crashes. Table 2 displays the descriptive statistics of significant attributes.

Table 1 shows that the distribution of injury severity among three time periods is quite similar. Overall, severe, moderate, possible, and no injuries constitute 6.3%, 22.6%, 20.0%, and 50.8%, of the total data. A few points are noteworthy in Table 2. First, some driver-related attributes such as seatbelt usage is found being statistically significant and impacting injury severity in all time periods. Also, this factor is overrepresented with more than 80% of the crash data. On the contrary, the factor of falling asleep while driving also turned out to significantly influence injury severity in all time periods. Still, it is underrepresented in crash data. This could be attributed to the difficulty of proving drowsiness as the main cause of accidents due to failures of states to come up with a standardized reporting system that can easily define sleepiness-related crashes. Female drivers and losing control of a vehicle while negotiating a horizontal curve account for more than 30% and 40% of the observations, respectively. Second, in terms of crash attributes, more than 70% of crashes along horizontal curves involved a single vehicle in all time periods.

5. Temporal stability tests

Likelihood ratio tests are commonly used to statistically verify whether developing separate models is justified or crash data can be aggregated to estimate a joint model. Herein, the temporal stability of determinants that impact the severity of drivers in RTU horizontal curves crashes can be examined by splitting crash data into three time periods: 2011–2012, 2013–2014, and 2015–2016. For details on two likelihood ratio tests, readers are referred to Washington et al. (2011). The test statistic of the first test is formulated as:

$$X^2 = -2[LL(\beta_{2011-2016}) - LL(\beta_{2011-2012}) - LL(\beta_{2013-2014}) - LL(\beta_{2015-2016})] \quad (4)$$

where $LL(\beta_{2011-2016})$ is the log-likelihood at convergence of the models estimated with whole crash data (i.e., for crashes occurred between 2011 and 2016). $LL(\beta_{2011-2012})$, $LL(\beta_{2013-2014})$, and$LL(\beta_{2015-2016})$ are the log-likelihood at convergence of the model using 2011–2012, 2013–2014, and 2015–2016 crash data, respectively. The outcome statistic X^2obtained using Eq. (4) equals to 1,747.58 with 41 degrees of freedom, which suggests that the null hypothesis that examining temporal stability by developing separate models for the specified time periods is inessential can be rejected with over 99.99% confidence. This finding confirms with considerable evidence that severity models regarding RTU horizontal curves crashes are temporally instable.

The second test is utilized to test the stability of injury severity determinants over time, this test can be written as in Eq. (5) (Washington et al., 2011):

$$X^2 = -2\left[LL\left(\beta_{t_2 t_1}\right) - LL(\beta_{t_1})\right] \quad (5)$$

The definition of each term is defined in Washington et al. (2011). According to the displayed results in Table 3, it is obvious that the null hypothesis that the parameters are equivalent in the two time periods can be rejected with more than 99.99% confidence level. The results of likelihood ratio tests collectively confirm that this study is in line with past studies in which the temporal instability of injury severity determinants was examined (Al-Bdairi et al., 2020; Alnawmasi & Mannering, 2019; Behnood & Al-Bdairi, 2020; Behnood & Mannering, 2016, 2015). However, these studies did not investigate crashes along horizontal curves.

Table 1
Distribution of injury severity and frequency.

Time period	Injury severity	Observations	Percent (%)
2011–2012	Severe	278	6.30%
	Moderate	1,009	22.87%
	Possible	883	20.01%
	No injury	2,242	50.82%
	Total	4,412	100.00%
2013–2014	Severe	292	6.39%
	Moderate	1,033	22.60%
	Possible	884	19.34%
	No injury	2,362	51.67%
	Total	4,571	100%
2015–2016	Severe	313	6.39%
	Moderate	1,099	22.43%
	Possible	1,018	20.78%
	No injury	2,469	50.40%
	Total	4,899	100.00%

6. Estimation results

The estimation results along with marginal effects for the developed models for 2011–2012, 2013–2014, and 2015–2016 time periods data, respectively, are presented in Tables 4–6. These tables reveal that heterogeneity in means and variances was captured in all three time period models. Also, Tables 4–6 show that all three estimated models have a good overall statistical fit as indicated by the values of McFadden Pseudo R-squared, which are in the range of 0.2–0.3. In this analysis, an extensive list of contributing factors that can significantly affect the injury severity outcomes of crashes along RTU horizontal curves is considered. However, only the variables with significant t-statistics with at least 90% significant level were kept in the models. By considering t-statistics as a threshold, 22, 23, and 26 factors are included in 2011–2012, 2013–2014, and 2015–2016 time period models, respectively. In these models, 3 parameters (crashes occurred on weekdays, drunk drivers, and belted drivers), 2 parameters (losing control of vehicle and belted drivers), and 5 parameters (young drivers-30 years and younger, losing control of vehicle, belted drivers, middle-aged drivers-between 30 and 65 years, and wet roadway surface condition) produced statistically significant random parameters with normal distribution in 2011–2012, 2013–2014,

[1] The split of dataset adopted in this study is considered after careful testing for any potential temporal instability in specified time periods. The criterion that should be followed in separating the data is maintaining a reasonable number of observations in each time period so that the accuracy and performance of the developed models cannot be affected. Followed this criterion which has been recommended by Ye and Lord (2014), it was found that splitting the data into 2011–2012, 2013–2014, and 2015–2016 time periods provided the only statistically significant separation.

Table 2
Summary statistics of significant attributes in the developed models.

Variable	2011–2012		2013–2014		2015–2016	
	Mean	S. D	Mean	S. D	Mean	S. D
Driver characteristics						
Elderly driver (1 if more than 65 years; 0 otherwise)	0.079	0.270	0.088	0.283	0.070	0.255
Drunk drivers (1 if drunk drivers; 0 otherwise)	0.125	0.331	0.123	0.328	0.106	0.308
Fatigued (1 if driver was fatigued; 0 otherwise)	–	–	–	–	0.051	0.221
Driver license status (1 if valid other states license; 0 otherwise)	0.071	0.257	–	–	–	–
Female (1 if driver is female; 0 otherwise)	0.360	0.480	0.367	0.482	0.286	0.452
Young driver (1 if less than 30 years; 0 otherwise)	–	–	–	–	0.367	0.482
Driver error (1 if driver failed to negotiate a curve; 0 otherwise)	–	–	0.045	0.206	0.104	0.305
Losing control of vehicle (1 if yes; 0 otherwise)	0.444	0.497	0.419	0.493	0.310	0.463
Belted drivers (1 if driver was belted; 0 otherwise)	0.832	0.374	0.823	0.381	0.649	0.477
Distracted (1 if driver was distracted; 0 otherwise)	–	–	0.036	0.186	0.037	0.189
Driver license status (1 if valid Oregon license; 0 otherwise)	–	–	–	–	0.638	0.481
Falling asleep (1 if yes; 0 otherwise)	0.054	0.226	0.050	0.218	0.049	0.215
Middle-aged driver (1 if between 30 and 65 years; 0 otherwise)	–	–	0.426	0.495	0.334	0.472
Roadway characteristics						
Wet (1 if wet roadway; 0 otherwise)	–	–	–	–	0.299	0.458
Speed limit (1 if 55 mph; 0 otherwise)	0.506	0.500	0.510	0.500	0.536	0.499
Dry (1 if dry roadway; 0 otherwise)	0.503	0.500	0.517	0.500	0.541	0.498
National highway system (1 if yes; 0 otherwise)	–	–	–	–	0.219	0.414
Speed limit (1 if 45 mph, 0 otherwise)	0.080	0.271	0.071	0.256	–	–
Curve sign (1 if traffic control device is curve sign; 0 otherwise)	0.033	0.179	0.039	0.194	–	–
Run-off (1 if yes; 0 otherwise)	0.892	0.310	0.895	0.307	0.874	0.332
No traffic control device (1 if yes; 0 otherwise)	–	–	–	–	0.336	0.472
Crash characteristics						
Single vehicle involved (1 if yes; 0 otherwise)	0.723	0.448	0.729	0.445	0.716	0.451
Head-on (1 if crash type is head-on; 0 otherwise)	0.083	0.275	0.075	0.264	0.082	0.275
Exceeding the posted speed limit (1 if yes; 0 otherwise)	0.039	0.192	0.039	0.195	0.055	0.227
Fixed object (1 if crash type is fixed object; 0 otherwise)	0.644	0.479	–	–	–	–
Airbag deployment (1 if airbag was deployed; 0 otherwise)	0.213	0.410	0.209	0.406	0.203	0.402
Overturned (1 if crash type is overturned; 0 otherwise)	0.080	0.272	0.061	0.239	0.073	0.260
Time related attributes						
Winter (1 if between January and April; 0 otherwise)	0.270	0.444	0.242	0.428	–	–
Morning (1 if between 4:00 AM and 11:00 AM; 0 otherwise)	–	–	–	–	0.235	0.424
Afternoon (1 if between 11:00 AM and 6:00 PM; 0 otherwise)	0.408	0.491	0.424	0.494	–	–
Night (1 if between 6:00 PM and 12:00 AM; 0 otherwise)	–	–	–	–	0.289	0.453
Weekdays (1 if yes; 0 otherwise)	0.658	0.474	–	–	–	–
Fall (1 if between September and December; 0 otherwise)	–	–	–	–	0.371	0.483
Weather and lighting attributes						
Dawn (1 if dawn lighting condition; 0 otherwise)	–	–	0.027	0.162	–	–
Clear weather (1 if yes; 0 otherwise)	0.487	0.500	–	–	–	–
Snowy weather (1 if yes; 0 otherwise)	–	–	0.042	0.200	–	–

S.D is standard deviation.

Table 3
Results of temporal stability tests with degrees of freedom for specified time periods (degrees of freedom in parentheses).

t_1	t_2		
	2011–2012	2013–2014	2015–2016
2011–2012	–	496.74 (23) [>99.99 %]	477.83 (26) [>99.99 %]
2013–2014	488.54 (22) [>99.99 %]	–	280.89 (26) [>99.99 %]
2015–2016	56.10 (22) [>99.99 %]	47.06 (23) [>99.78 %]	–

Note: values in brackets are confidence level.

and 2015–2016 time period models, respectively. Thus, to ease interpretation of the findings due to the vast number of explanatory variables being statistically significant in each model, these variables were categorized into driver characteristics, roadway characteristics, crash characteristics, time-related attributes, and weather and lighting attributes. As seen in Tables 4–6, some factors were found to affect exclusively one time period while others turned out to be statistically significant in all or at least two time period models. Accordingly, to highlight the differences in their influence on injury severity, the marginal effects of contributing factors in all time period models will be presented in Table 7 side-by-side. The findings of the current study will be elaborated in more details in the next subsections.

6.1. Driver characteristics

The estimation results disclose that some driver-related attributes affect injury severity of crashes at RTU horizontal curves in all time period models. However, these factors are temporally instable, and their impacts are different (as shown in Table 7). Specifically, six parameters produced statistically significant impacts on injury severity outcomes in all specified time period models. These parameters include older drivers with more than 65 years, drunk drivers, female drivers, belted drivers, losing control of the vehicle, and driver fell asleep. The indicator variable of belted drivers was found to produce random parameter in the three models with overall increase in the likelihood of moderate injury. The values of mean (standard deviation) of this indicator variable were obtained as -1.551 (1.502), -3.696 (4.565), and -12.822 (13.907) in 2011–2012, 2013–2014, and 2015–2016 time period models, respectively. This implies that moderate injury likelihood will be more likely to be sustained by belted drivers for

Table 4
Estimation results for 2011–2012 time period model.

Variable	Parameter estimate	t-stat	Marginal effects			
			Severe injury	Moderate injury	Possible injury	No injury
Constant [SI]	−4.299	−14.86	–	–	–	–
Constant [MI]	−2.699	−10.46	–	–	–	–
Constant [PI]	−2.241	−12.92	–	–	–	–
Driver characteristics						
Elderly driver (1 if more than 65 years; 0 otherwise) [SI]	0.806	3.66	**0.0045**	−0.0007	−0.0009	−0.0028
Drunk drivers (1 if drunk drivers; 0 otherwise) [SI]	1.113	2.28	**0.0254**	−0.0059	−0.0051	−0.0144
Standard deviation of "drunk drivers" (normally distributed)	*1.628*	*2.10*	–	–	–	–
Female (1 if driver is female; 0 otherwise) [PI]	0.506	3.58	−0.0016	−0.0042	**0.0204**	−0.0146
Falling asleep (1 if yes; 0 otherwise) [NI]	−0.534	−2.95	0.0010	0.0019	0.0022	**−0.0052**
Losing control of vehicle (1 if yes; 0 otherwise) [MI]	0.436	3.38	−0.0016	**0.0211**	−0.0047	−0.0148
Belted driver (1 if driver was belted; 0 otherwise) [MI]	−1.551	−2.55	0.0059	**−0.0584**	0.0190	0.0334
Standard deviation of "belted driver" (normally distributed)	*1.502*	*2.38*	–	–	–	–
Driver license status (1 if valid other states license; 0 otherwise) [NI]	0.264	1.71	−0.0005	−0.0013	−0.0014	**0.0032**
Roadway characteristics						
Curve sign (1 if traffic control device is curve sign; 0 otherwise) [SI]	−0.773	−2.38	**−0.0020**	0.0013	0.0003	0.0004
Dry (1 if dry roadway; 0 otherwise) [MI]	0.964	5.85	−0.0051	**0.0536**	−0.0122	−0.0363
Speed limit (1 if 55 mph, 0 otherwise) [MI]	0.547	4.11	−0.0024	**0.0290**	−0.0064	−0.0202
Run-off (1 if yes; 0 otherwise) [NI]	−0.552	−3.49	0.0152	0.0343	0.0390	**−0.0885**
Speed limit (1 if 45 mph; 0 otherwise) [NI]	−0.398	−2.74	0.0012	0.0019	0.0026	**−0.0056**
Crash characteristics						
Single vehicle involved (1 if yes; 0 otherwise) [SI]	0.663	2.63	**0.0230**	−0.0041	−0.0044	−0.0145
Exceeding the posted speed limit (1 if yes; 0 otherwise) [SI]	0.782	2.75	**0.0026**	−0.0005	−0.0005	−0.0016
Fixed object (1 if crash type is fixed object; 0 otherwise) [PI]	0.515	3.91	−0.0030	−0.0080	**0.0377**	−0.0268
Overturned (1 if crash type is overturned; 0 otherwise) [NI]	−0.636	−4.18	0.0016	0.0045	0.0036	**−0.0097**
Head-on (1 if collision type is head-on; 0 otherwise) [SI]	2.072	7.05	**0.0174**	−0.0027	−0.0030	−0.0117
Airbag deployment (1 if airbag was deployed; 0 otherwise) [SI]	1.330	8.02	**0.0226**	−0.0030	−0.0043	−0.0153
Time related attributes						
Winter (1 if between January and April; 0 otherwise) [SI]	−0.520	−2.71	**−0.0046**	0.0007	0.0008	0.0031
Afternoon (1 if between 11:00 AM and 6:00 PM; 0 otherwise) [MI]	−0.233	−1.87	0.0006	**−0.0086**	0.0019	0.0061
Weekdays (1 if yes; 0 otherwise) [PI]	−0.877	−1.24	−0.0048	−0.0153	**0.0478**	−0.0277
Standard deviation of "weekdays" (normally distributed)	*1.673*	*1.73*	–	–	–	–
Weather and lighting attributes						
Clear weather (1 if yes; 0 otherwise) [SI]	0.318	2.12	**0.0082**	−0.0016	−0.0015	−0.0052
Heterogeneity in the means of the random parameters						
Belted driver: Single vehicle involved (1 if yes; 0 otherwise)	1.251	3.85	–	–	–	–
Belted driver: Head-on (1 if crash type is head-on; 0 otherwise)	1.610	3.95	–	–	–	–
Weekdays: Head-on (1 if crash type is head-on; 0 otherwise)	1.322	3.62	–	–	–	–
Weekdays: Elderly driver (1 if more than 65 years; 0 otherwise)	0.568	1.92	–	–	–	–
Weekdays: Dry (1 if dry roadway; 0 otherwise)	0.384	2.31	–	–	–	–
Heterogeneity in the variances of the random parameters						
Belted driver: Airbag deployment (1 if airbag was deployed; 0 otherwise)	1.391	5.37	–	–	–	–
Belted driver: Female (1 if driver is female; 0 otherwise)	0.435	2.42	–	–	–	–
Weekdays: Airbag deployment (1 if airbag was deployed; 0 otherwise)	0.658	2.34	–	–	–	–
Weekdays: Female (1 if driver is female; 0 otherwise)	0.435	2.46	–	–	–	–
Model statistics						
Number of observations				4412		
Log likelihood at convergence				−4818.78		
Log likelihood at constants only				−6116.33		
McFadden Pseudo R-squared				0.212		

Note: letters in brackets are explanatory variable defined for: [NI] No injury; [PI] Possible injury; [MI] Moderate injury; [SI]: Severe injury.

15.1%, 20.1%, and 17.8% of the observations in 2011–2012, 2013–2014, and 2015–2016 time period models, respectively. Looking into these percentages, one could conclude that there is no substantial difference regarding how moderate injury outcome increases as time passes. Fig. 1 shows that marginal effects of belted drivers corresponding to moderate injury levels are decreased meaning that the probability of ending up with moderate injuries in crashes occurred along RTU horizontal curves in which drivers are belted are highly unlikely, while no injury outcomes would be more likely to be sustained by those drivers. This finding is intuitive and underscores the importance of safety seatbelt in saving lives and reducing injuries resulted from roadway crashes. Previous studies also documented the significance of seatbelt use and how it reduces the chances of fatal crashes (Russo et al., 2014; Schneider et al., 2009).

Losing control of the vehicle was found to produce statistically significant random parameter in two time period models (i.e., 2013–2014 and 2015–2016), while this variable was found to be fixed in 2011–2012 model. For random parameters, the means (standard deviations) are −0.849 (2.849) and −7.014 (19.879) for 2013–2014 and 2015–2016 time period models, respectively. This distribution implies that 38.3% and 36.2% of crashes occurred due to losing control of the vehicle will increase the probability of sustaining moderate injuries in crashes along RTU horizontal curves for 2013–2014 and 2015–2016 time period models, respectively. As shown in Fig. 2, the moderate injury outcomes are more likely

Table 5
Estimation results for 2013–2014 time period model.

Variable	Parameter estimate	t-stat	Marginal effects			
			Severe injury	Moderate injury	Possible injury	No injury
Constant [SI]	−4.034	−15.65	–	–	–	–
Constant [MI]	−2.496	−9.43	–	–	–	–
Constant [PI]	−1.672	−13.07	–	–	–	–
Driver characteristics						
Elderly driver (1 if more than 65 years; 0 otherwise) [SI]	1.043	5.22	**0.0081**	−0.0009	−0.0019	−0.0053
Drunk drivers (1 if drunk drivers; 0 otherwise) [SI]	1.843	12.32	**0.0304**	−0.0049	−0.0067	−0.0189
Losing control of vehicle (1 if yes; 0 otherwise) [MI]	−0.849	−1.68	−0.0016	**0.0112**	−0.0024	−0.0072
Standard deviation of "losing control of vehicle" (normally distributed)	*2.849*	*3.08*	–	–	–	–
Failing to negotiate curve (1 if yes; 0 otherwise) [SI]	0.664	2.73	**0.0028**	−0.0003	−0.0007	−0.0018
Belted driver (1 if driver was belted; 0 otherwise) [MI]	−3.696	−2.29	0.0069	**−0.0686**	0.0183	0.0434
Standard deviation of "belted driver" (normally distributed)	*4.565*	*3.01*	–	–	–	–
Female (1 if driver is female; 0 otherwise) [PI]	0.759	9.48	−0.0041	−0.0061	**0.0484**	−0.0381
Distracted (1 if driver was distracted; 0 otherwise) [PI]	0.450	2.24	−0.0002	−0.0004	**0.0028**	−0.0022
Falling asleep (1 if yes; 0 otherwise) [NI]	−0.603	−3.54	0.0013	0.0012	0.0032	**−0.0058**
Middle-aged driver (1 if between 30 and 65 years; 0 otherwise) [NI]	−0.240	−3.24	0.0034	0.0041	0.0122	**−0.0197**
Roadway characteristics						
Dry (1 if dry roadway; 0 otherwise) [MI]	0.524	2.65	−0.0020	**0.0186**	−0.0047	−0.0118
Speed limit (1 if 55 mph, 0 otherwise) [MI]	0.824	3.83	−0.0028	**0.0285**	−0.0072	−0.0184
Run-off (1 if yes; 0 otherwise) [NI]	−0.241	−2.02	0.0075	0.0089	0.0248	**−0.0412**
Speed limit (1 if speed limit is 45 mph; 0 otherwise) [NI]	−0.321	−2.33	0.0009	0.0007	0.0027	**−0.0042**
Curve sign (1 if traffic control device is curve sign; 0 otherwise) [MI]	1.081	2.57	−0.0003	**0.0032**	−0.0008	−0.0021
Crash characteristics						
Single vehicle involved (1 if yes; 0 otherwise) [SI]	0.807	3.47	**0.0330**	−0.0039	−0.0081	−0.0210
Head-on (1 if crash type is head-on; 0 otherwise) [SI]	1.658	6.02	**0.0127**	−0.0014	−0.0028	−0.0085
Overturned (1 if crash type is overturned; 0 otherwise) [NI]	−0.438	−2.88	0.0009	0.0011	0.0032	**−0.0052**
Exceeding the posted speed limit (1 if yes; 0 otherwise) [SI]	0.554	1.99	**0.0018**	−0.0002	−0.0004	−0.0012
Airbag deployment (1 if airbag was deployed; 0 otherwise) [SI]	1.044	7.36	**0.0215**	−0.0026	−0.0053	−0.0136
Time related attributes						
Winter (1 if between January and April; 0 otherwise) [SI]	−0.610	−3.44	**−0.0054**	0.0005	0.0013	0.0036
Afternoon (1 if between 11:00 AM and 6:00 PM; 0 otherwise) [MI]	−0.505	−2.50	0.0010	**−0.0125**	0.0032	0.0083
Weather and lighting attributes						
Dawn (1 if dawn lighting condition; 0 otherwise) [SI]	0.744	2.21	**0.0017**	−0.0002	−0.0004	−0.0011
Snowy weather (1 if yes; 0 otherwise) [PI]	−0.686	−2.99	0.0002	0.0002	**−0.0028**	0.0024
Heterogeneity in means of random parameters						
Losing control of vehicle: Drunk drivers (1 if drunk drivers; 0 otherwise)	1.000	2.24	–	–	–	–
Losing control of vehicle: Dawn (1 if dawn lighting condition; 0 otherwise)	−2.817	−2.05	–	–	–	–
Belted driver: Single vehicle involved (1 if yes; 0 otherwise)	2.246	3.01	–	–	–	–
Belted driver: Female (1 if driver is female; 0 otherwise)	1.079	3.24	–	–	–	–
Belted driver: Head-on (1 if crash type is head-on; 0 otherwise)	2.189	2.90	–	–	–	–
Belted driver: Winter (1 if between January and April; 0 otherwise)	−0.906	−2.48	–	–	–	–
Belted driver: Drunk drivers (1 if drunk drivers; 0 otherwise)	1.307	2.28	–	–	–	–
Heterogeneity in variances of random parameters						
Belted driver: Speed limit (1 if 45 mph, 0 otherwise)	0.350	2.15	–	–	–	–
Model statistics						
Number of observations				4571		
Log likelihood at convergence				−5024.85		
Log likelihood at constants only				−6336.75		
McFadden Pseudo R-squared				0.207		

Note: letters in brackets are explanatory variable defined for: [NI] No injury; [PI] Possible injury; [MI] Moderate injury; [SI]: Severe injury.

to be incurred by drivers losing control of their vehicles when involved in crashes along RTU horizontal curves. Also, this figure shows that the trend regarding the effect of this variable is declining over time. The possible reason could be due to the drivers' behavior related to negotiating curves by compensating that with reducing speed, which in turn alleviates the severity of crash outcomes. This growing attitude among drivers could be due to unstopping efforts lead by safety officials and traffic engineers to enhance safety at such hazardous locations.

Driver age turned out to highly affect the outcome of injury severity in crashes along RTU horizontal curves. For example, older drivers (more than 65 years) are more prone to severe injury levels compared to other drivers when they are involved in crashes along RTU horizontal curves, regardless of the time period used. Table 7 and Fig. 3 disclose that the probability of severe injury levels for older drivers is nearly 6.4, 9.0, and 37.5 times compared to moderate injuries in 2011–2012, 2013–2014, and 2015–2016 time period models, respectively. This finding explains how older drivers who encompass large share from the U.S. drivers are at high risk, particularly at horizontal curves. A possible reason could be related to deterioration of driving abilities of older drivers and reaction time in negotiating curves. Also, the increased chances of injury levels being sustained by such drivers are attributed to fragility of older drivers' bodies compared to young and middle-aged drivers. Similar findings were obtained in past studies (Kim et al., 2013; Schneider et al., 2009).

Alcohol consumption has been extensively documented as a leading cause of crashes because such behavior has negative impli-

Table 6
Estimation results for 2015–2016 time period model.

Variable	Parameter estimate	t-stat	Marginal effects			
			Severe injury	Moderate injury	Possible injury	No injury
Constant [SI]	−5.422	−14.17	–	–	–	–
Constant [MI]	−3.844	−12.87	–	–	–	–
Constant [PI]	−3.796	−14.27	–	–	–	–
Driver characteristics						
Drunk drivers (1 if drunk drivers; 0 otherwise) [SI]	2.136	11.93	**0.0244**	−0.0017	−0.0073	−0.0155
Failing to negotiate curve (1 if yes; 0 otherwise) [SI]	1.066	5.45	**0.0085**	−0.0004	−0.0030	−0.0050
Losing control of vehicle (1 if yes; 0 otherwise) [MI]	−7.014	−1.46	−0.0009	**0.0119**	−0.0050	−0.0060
Standard deviation of "losing control of vehicle" (normally distributed)	*19.879*	*2.00*	–	–	–	–
Elderly driver (1 if more than 65 years; 0 otherwise) [SI]	1.362	6.12	**0.0075**	−0.0002	−0.0017	−0.0056
Fatigued (1 if driver was fatigued; 0 otherwise) [SI]	−0.875	−2.72	**−0.0021**	0.0001	0.0009	0.0011
Belted driver (1 if driver was belted; 0 otherwise) [MI]	−12.822	−2.01	0.0015	**−0.0198**	0.0090	0.0093
Standard deviation of "belted driver" (normally distributed)	*13.907*	*2.49*	–	–	–	–
Distracted (1 if driver was distracted; 0 otherwise) [MI]	2.099	4.89	−0.0004	**0.0034**	−0.0012	−0.0018
Female (1 if driver is female; 0 otherwise) [PI]	1.207	8.91	−0.0059	−0.0026	**0.0351**	−0.0266
Young driver (1 if less than 30 years; 0 otherwise) [PI]	−1.656	−2.19	−0.0018	−0.0069	**0.0302**	−0.0215
Standard deviation of "young driver" (normally distributed)	*3.587*	*4.30*	–	–	–	–
Driver license status (1 if valid Oregon license; 0 otherwise) [PI]	1.721	9.46	−0.0177	−0.0072	**0.1046**	−0.0797
Falling asleep (1 if yes; 0 otherwise) [NI]	−1.322	−5.31	0.0019	0.0007	0.0037	**−0.0063**
Middle-aged driver (1 if between 30 and 65 years; 0 otherwise) [NI]	−2.040	−7.27	0.0115	0.0030	0.0356	**−0.0501**
Standard deviation of "middle-aged driver" (normally distributed)	*1.413*	*3.23*	–	–	–	–
Roadway characteristics						
National highway system (1 if yes; 0 otherwise) [PI]	0.481	3.45	−0.0015	−0.0005	**0.0088**	−0.0068
Wet (1 if wet roadway; 0 otherwise) [SI]	−1.315	−2.18	**0.0043**	−0.0002	−0.0011	−0.0030
Standard deviation of "wet" (normally distributed)	*1.665*	*3.39*	–	–	–	–
Speed limit (1 if 55 mph; 0 otherwise) [SI]	0.586	3.93	**0.0154**	−0.0007	−0.0050	−0.0098
Dry (1 if dry roadway; 0 otherwise) [MI]	0.484	1.88	−0.0007	**0.0067**	−0.0016	−0.0045
No traffic control device (1 if yes; 0 otherwise) [NI]	0.226	1.90	−0.0018	−0.0011	−0.0046	**0.0075**
Run-off (1 if yes; 0 otherwise) [NI]	−0.491	−2.63	0.0115	0.0067	0.0252	**−0.0433**
Crash characteristics						
Single vehicle involved (1 if yes; 0 otherwise) [SI]	1.103	3.54	**0.0342**	−0.0016	−0.0108	−0.0218
Head-on (1 if crash type is head-on; 0 otherwise) [SI]	2.220	6.67	**0.0154**	−0.0006	−0.0046	−0.0103
Overturned (1 if crash type is overturned; 0 otherwise) [NI]	−0.350	−1.92	0.0007	0.0003	0.0013	**−0.0024**
Exceeding the posted speed limit (1 if yes; 0 otherwise) [SI]	0.918	3.49	**0.0033**	−0.0001	−0.0008	−0.0023
Airbag deployment (1 if airbag was deployed; 0 otherwise) [SI]	1.303	7.75	**0.0170**	−0.0005	−0.0056	−0.0109
Time related attributes						
Morning (1 if between 4:00 AM and 11:00 AM; 0 otherwise) [SI]	0.363	2.06	**0.0036**	−0.0001	−0.0011	−0.0024
Night (1 if between 6:00 PM and 12:00 AM; 0 otherwise) [MI]	0.564	2.23	−0.0004	**0.0045**	−0.0010	−0.0031
Fall (1 if between September and December; 0 otherwise) [NI]	0.310	2.60	−0.0028	−0.0016	−0.0069	**0.0112**
Heterogeneity in the means of the random parameters						
Wet: Single vehicle involved (1 if yes; 0 otherwise)	−0.883	−2.00	–	–	–	–
Wet: Driver license status (1 if valid Oregon license; 0 otherwise)	1.333	2.94	–	–	–	–
Belted driver: Single vehicle involved (1 if yes; 0 otherwise)	7.867	2.32	–	–	–	–
Young driver: Single vehicle involved (1 if yes; 0 otherwise)	1.835	3.27	–	–	–	–
Middle-aged driver: Single vehicle involved (1 if yes; 0 otherwise)	−0.425	−2.11	–	–	–	–
Middle-aged driver: Driver license status (1 if valid Oregon license; 0 otherwise)	1.159	4.38	–	–	–	–
Heterogeneity in the variances of the random parameters						
Losing control of vehicle: Airbag deployment (1 if airbag was deployed; 0 otherwise)	1.877	4.91	–	–	–	–
Belted driver: Overturned (1 if crash type is overturned; 0 otherwise)	1.260	3.67	–	–	–	–
Belted driver: Airbag deployment (1 if airbag was deployed; 0 otherwise)	1.254	6.24	–	–	–	–
Young driver: Airbag deployment (1 if airbag was deployed; 0 otherwise)	1.480	3.75	–	–	–	–
Model statistics						
Number of observations				4899		
Log likelihood at convergence				−4726.49		
Log likelihood at constants only				−6791.46		
McFadden Pseudo R-squared				0.304		

Note: letters in brackets are explanatory variable defined for: [NI] No injury; [PI] Possible injury; [MI] Moderate injury; [SI]: Severe injury.

cations on drivers' awareness, consciousness, reaction time, and driving abilities. Tables 4–6 clearly show that drunk drivers are more prone to severe injuries when involved in crashes along horizontal curves. Table 7 and Fig. 4 illustrate that the probability of being involved in severe injuries among drunk drivers is 4.3, 6.2, and 14.4 times in comparison with moderate injuries in 2011–2012, 2013–2014, and 2015–2016 time period models, respectively. Based on this finding, strict law enforcement on alcohol consumption is needed to alleviate this behavior.

6.2. Roadway characteristics

Roadway surface condition is an important determinant of injury severity. In this study, dry roadway surface condition, as mentioned above, was found to increase the likelihood of moderate

Table 7
Marginal effects of injury severity in different time period models.

Variable	Severe injury			Moderate injury			Possible injury			No injury		
	2011–12	2013–14	2015–16	2011–12	2013–14	2015–16	2011–12	2013–14	2015–16	2011–12	2013–14	2015–16
Driver characteristics												
Elderly driver (1 if more than 65 years; 0 otherwise)	0.0045	0.0081	0.0075	−0.0007	−0.0009	−0.0002	−0.0009	−0.0019	−0.0017	−0.0028	−0.0053	−0.0056
Drunk drivers (1 if drunk drivers; 0 otherwise)	0.0254	0.0304	0.0244	−0.0059	−0.0049	−0.0017	−0.0051	−0.0067	−0.0073	−0.0144	−0.0189	−0.0155
Fatigued (1 if driver was fatigued; 0 otherwise)	–	–	−0.0021	–	–	0.0001	–	–	0.0009	–	–	0.0011
Driver license status (1 if valid other states license; 0 otherwise)	−0.0005	–	–	−0.0013	–	–	−0.0014	–	–	0.0032	–	–
Female (1 if driver is female; 0 otherwise)	−0.0016	−0.0041	−0.0059	−0.0042	−0.0061	−0.0026	0.0204	0.0484	0.0351	−0.0146	−0.0381	−0.0266
Young driver (1 if less than 30 years; 0 otherwise)	–	–	−0.0018	–	–	−0.0069	–	–	0.0302	–	–	−0.0215
Failing to negotiate curve (1 if yes; 0 otherwise)	–	0.0028	0.0085	–	−0.0003	−0.0004	–	−0.0007	−0.0030	–	−0.0018	−0.0050
Losing control of vehicle (1 if yes; 0 otherwise)	−0.0016	−0.0016	−0.0009	0.0211	0.0112	0.0119	−0.0047	−0.0024	−0.0050	−0.0148	−0.0072	−0.0060
Belted drivers (1 if driver was belted; 0 otherwise)	0.0059	0.0069	0.0015	−0.0584	−0.0686	−0.0198	0.0190	0.0183	0.0090	0.0334	0.0434	0.0093
Distracted (1 if driver was distracted; 0 otherwise)	–	−0.0002	−0.0004	–	−0.0004	0.0034	–	0.0028	−0.0012	–	−0.0022	−0.0018
Driver license status (1 if valid Oregon license; 0 otherwise)	–	–	−0.0177	–	–	−0.0072	–	–	0.1046	–	–	−0.0797
Falling asleep (1 if yes; 0 otherwise)	0.0010	0.0013	0.0019	0.0019	0.0012	0.0007	0.0022	0.0032	0.0037	−0.0052	−0.0058	−0.0063
Middle-aged driver (1 if between 30 and 65 years; 0 otherwise)	–	0.0034	0.0115	–	0.0041	0.0030	–	0.0122	0.0356	–	−0.0197	−0.0501
Roadway characteristics												
Wet (1 if wet roadway; 0 otherwise)	–	–	0.0043	–	–	−0.0002	–	–	−0.0011	–	–	−0.0030
Speed limit (1 if 55 mph; 0 otherwise)	−0.0024	−0.0028	0.0154	0.0290	0.0285	−0.0007	−0.0064	−0.0072	−0.0050	−0.0202	−0.0184	−0.0098
Dry (1 if dry roadway; 0 otherwise)	−0.0051	−0.0020	−0.0007	0.0536	0.0186	0.0067	−0.0122	−0.0047	−0.0016	−0.0363	−0.0118	−0.0045
National highway system (1 if yes; 0 otherwise)	–	–	−0.0015	–	–	−0.0005	–	–	0.0088	–	–	−0.0068
Speed limit (1 if 45 mph; 0 otherwise)	0.0012	0.0009	–	0.0019	0.0007	–	0.0026	0.0027	–	−0.0056	−0.0042	–
Curve sign (1 if traffic control device is curve sign; 0 otherwise)	−0.0020	−0.0003	–	0.0013	0.0032	–	0.0003	−0.0008	–	0.0004	−0.0021	–
Run-off (1 yes; 0 otherwise)	0.0152	0.0075	0.0115	0.0343	0.0089	0.0067	0.0390	0.0248	0.0252	−0.0885	−0.0412	−0.0433
No traffic control device (1 yes; 0 otherwise)	–	–	−0.0018	–	-	−0.0011	–	–	−0.0046	–	–	0.0075
Crash characteristics												
Single vehicle involved (1 if yes; 0 otherwise)	0.0230	0.0330	0.0342	−0.0041	−0.0039	−0.0016	−0.0044	−0.0081	−0.0108	−0.0145	−0.0210	−0.0218
Head-on (1 if crash type is head-on; 0 otherwise)	0.0174	0.0127	0.0154	−0.0027	−0.0014	−0.0006	−0.0030	−0.0028	−0.0046	−0.0117	−0.0085	−0.0103
Exceeding the posted speed limit (1 yes; 0 otherwise)	0.0026	0.0018	0.0033	−0.0005	−0.0002	−0.0001	−0.0005	−0.0004	−0.0008	−0.0016	−0.0012	−0.0023
Fixed object (1 if crash type is fi-ed object; 0 otherwise)	−0.0030	–	–	−0.0080	–	–	0.0377	–	–	−0.0268	–	–
Airbag deployment (1 if airbag was deployed; 0 otherwise)	0.0226	0.0215	0.0170	−0.0030	−0.0026	−0.0005	−0.0043	−0.0053	−0.0056	−0.0153	−0.0136	−0.0109
Overturned (1 if crash type is overturned; 0 otherwise)	0.0016	0.0009	0.0007	0.0045	0.0011	0.0003	0.0036	0.0032	0.0013	−0.0097	−0.0052	−0.0024
Time related attributes												
Winter (1 if between January and April; 0 otherwise)	−0.0046	−0.0054	–	0.0007	0.0005	–	0.0008	0.0013	–	0.0031	0.0036	–
Morning (1 if between 4:00 AM and 11:00 AM; 0 otherwise)	–	–	0.0036	–	–	−0.0001	–	–	−0.0011	–	–	−0.0024
Afternoon (1 if between 11:00 AM and 6:00 PM; 0 otherwise)	0.0006	0.001	–	−0.0086	−0.0125	–	0.0019	0.0032	–	0.0061	0.0083	–
Night (1 if between 6:00 PM and 12:00 AM; 0 otherwise)	–	–	−0.0004	–	–	0.0045	–	–	−0.0010	–	–	−0.0031

Table 7 (*continued*)

Variable	Severe injury			Moderate injury			Possible injury			No injury		
	2011–12	2013–14	2015–16	2011–12	2013–14	2015–16	2011–12	2013–14	2015–16	2011–12	2013–14	2015–16
Weekdays (1 if yes; 0 otherwise)	−0.0048	–	–	−0.0153	–	–	0.0478	–	–	−0.0277	–	–
Fall (1 if between September and December; 0 otherwise)	–	–	−0.0028	–	–	−0.0016	–	–	−0.0069	–	–	0.0112
Weather and lighting attributes												
Dawn (1 if dawn lighting condition; 0 otherwise)	–	0.0017	–	–	−0.0002	–	–	−0.0004	–	–	−0.0011	–
Clear weather (1 if yes; 0 otherwise)	0.0082	–	–	−0.0016	–	–	−0.0015	–	–	−0.0052	–	–
Snowy weather (1 if yes; 0 otherwise)	–	0.0002	–	–	0.0002	–	–	−0.0028	–	–	0.0024	–

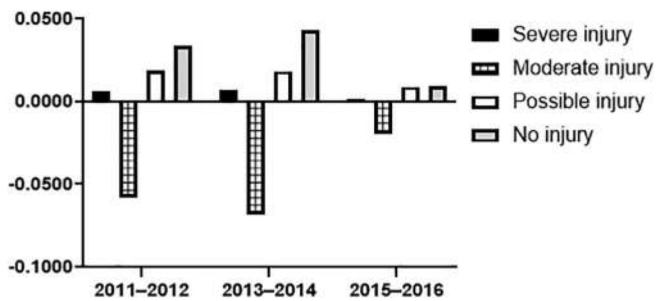


Fig. 1. Marginal effects of the indicator variable of belted drivers.

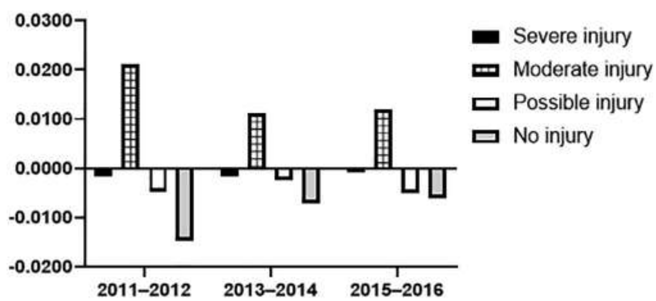


Fig. 2. Marginal effects of the indicator variable of losing control of vehicle.

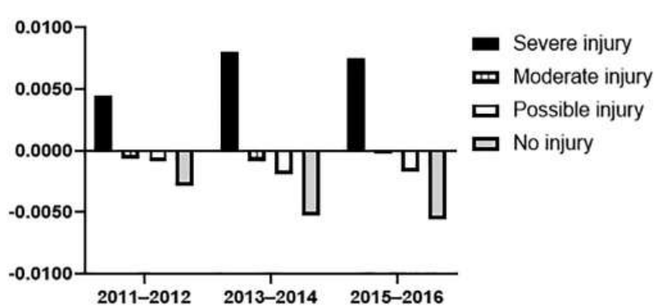


Fig. 3. Marginal effects of the indicator variable of older drivers.

injury severity in all time period models. Again, the overconfidence of drivers in such conditions and their attempts to increase their driving speed could be a reason. Unlike dry roadway surface condition, the propensity of severe injury levels would be higher under wet roadway surfaces. Table 6 reveals that in 2015–2016 time period model, crashes that happened along horizontal curves under wet roadways are more likely to end up with severe injuries. Fig. 5 also provides a clearer picture of how driving under such conditions jeopardizes drivers' safety as the probability of sustaining severe injuries is about 21.5 times compared to moderate injuries.

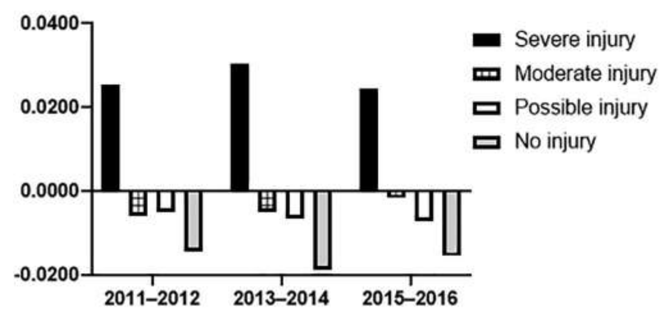


Fig. 4. Marginal effects of the indicator variable of drunk drivers.

Another notable factor that was found to significantly affect injury severity is traffic control device. For example, the presence of a curve sign as a traffic control device in the horizontal curves showed high instability in terms of its benefits on reducing crashes and injuries of these crashes at horizontal curves, because such a factor was found to statistically decrease the possibility of severe injuries in the 2011–2012 time period model, while the same factor highly increases the likelihood of moderate injuries in 2013–2014 model (as shown in Fig. 6). This variation or instability of this factor can translate the disparity of drivers in terms of their compliance with traffic control devices at horizontal curves. Contrarily, no injury outcomes are more likely being suffered in crashes occurring at horizontal curves in which no traffic control devices are used (see 2015–2016 time period model).

6.3. Crash characteristics

Five factors produced statistically significant impacts on injury severity in all time period models. It is noteworthy that these factors tend to be stable over time in their impacts on injury severity probabilities. These factors include single vehicle involved, head-on collision, overturn crashes, deployed airbags, and exceeding

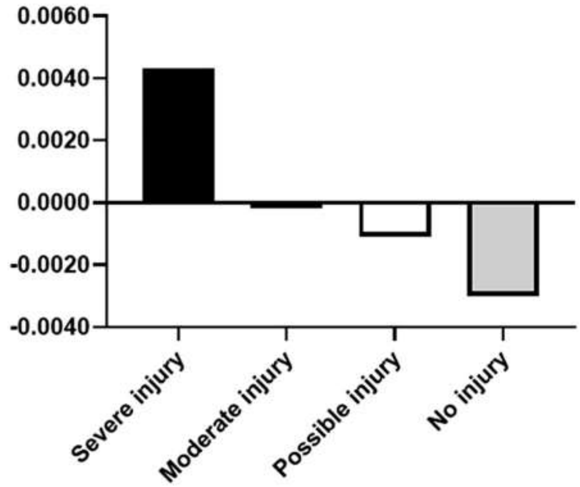


Fig. 5. Marginal effects of the indicator variable of wet roadway surface condition.

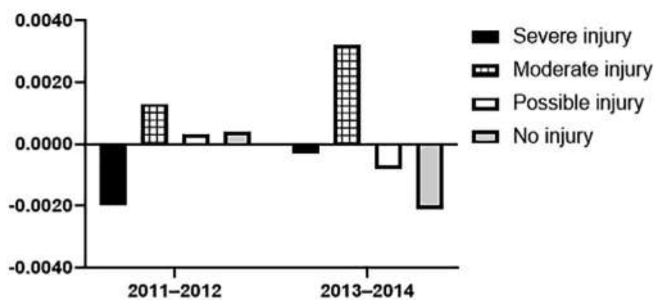


Fig. 6. Marginal effects of the indicator variable of curve sign.

the posted speed limit as the primary cause of the crash. Regarding the number of vehicles involved in crashes along RTU horizontal curves, crashes involving a single vehicle were found to significantly increase the propensity of severe injuries in all time periods, meaning that the effect of this variable is stable over time (as illustrated in Fig. 7). Driver errors such as failure of lane keeping, inability to identify the change in horizontal alignment, and speed while negotiating the curves are the most possible reasons underlying higher severity sustained in crashes along RTU horizontal curves involving a single vehicle compared to crashes involving more than one vehicle.

Crash type was found to be significant in terms of its effects on injury severity sustained. In this study, two crash types were identified to be dominant at RTU horizontal curves, namely head-on and overturn collisions. Both crashes affect injury severity outcomes in all time period models. The indicator variable of head-on crashes was found to increase the probability of severe injuries in all specified models, while the indicator variable of overturn crashes was found to be associated with less severe injuries (as seen in Table 7 and Fig. 8). This finding simply indicates how solid the problem is at RTU horizontal curves regarding head-on collisions over time. As such, proper and effective engineering measures that deter such crashes at RTU horizontal curves are urgently required. This finding is in complete agreement with previous studies (Anarkooli & Hosseinlou, 2016; Kockelman & Kweon, 2002; Lee & Li, 2014).

Among risky driving behavior factors that have direct implications on roadway safety is exceeding the speed limit. This variable was found to highly impact the probability of injury severities, particularly increases the severe injury severity in all time period models. From the findings, this variable shows relative stability

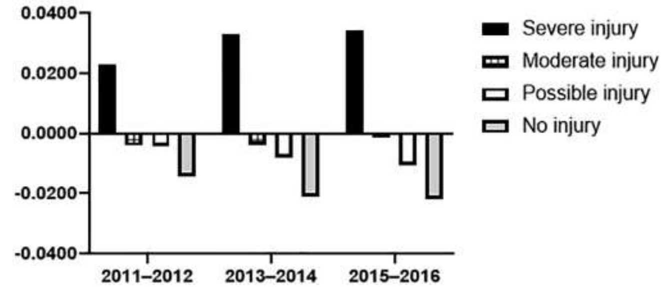


Fig. 7. Marginal effects of the indicator variable of single vehicle involved.

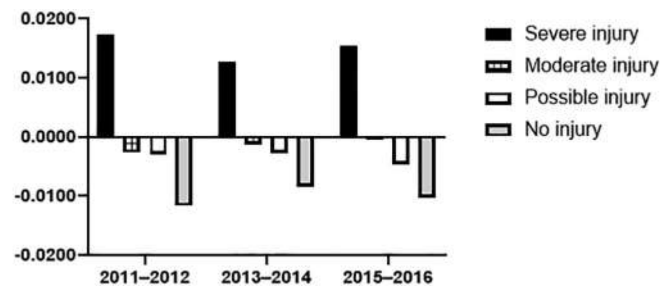


Fig. 8. Marginal effects of the indicator variable of head-on collisions.

over time. Yet, the marginal effects (see Table 7 and Fig. 9), illustrate that the influence of this variable increases in the last time period model. The reason underlying increased severity of crashes caused by exceeding the speed limit is attributed to the decreased perception-reaction time at a speed higher than the posted one.

6.4. Time related attributes

As documented in previous studies (Al-Bdairi, 2020; Behnood & Mannering, 2019; Pahukula et al., 2015; Zou et al., 2017), time of day is a significant attribute that can influence the injury severity of roadway crashes. In this research, three time of day periods produced significant effects on injury severity. These time of day periods are crashes occurring in morning (between 4 a.m. and 11 a.m.), crashes occurring in the afternoon (between 11 a.m. and 6p.m.), and crashes at night (between 6 p.m. and 12 a.m.). The first time of day period was only significant in the 2015–2016 time period model, the second time of day period was significant in the 2011–2012 and 2013–2014 time period models, and the third time of day period was only significant in the 2015–2016 time period model. For crashes that occurred at afternoon time of day, moderate injuries are highly unlikely to be sustained in both the 2011–2012 and the 2013–2014 time period models (see Fig. 10). This is attributed to rush hours at such time of day and the increase in traffic volumes, which reduce the propensity of higher injury levels.

In the 2015–2016 time period model, two time of day periods were found to affect the injury severity in crashes along RTU horizontal curves, morning and night. The first factor increases the severity of crashes conditioning that crashes occur at horizontal curves, while the effect of the latter factor is less compared to crashes occurring in the morning (as illustrated in Fig. 11). The disparity in effects of these time of day periods on injury severity is related to driver fatigue and the effect of alcohol being consumed on the performance of drivers in the morning (between 4 a.m. and 11 a.m.) time window that may increase the chance of drivers being involved in severe crashes. In contrast, at night (between 6 p.m. and 12:00 a.m.) crashes are relatively less severe due to the fact

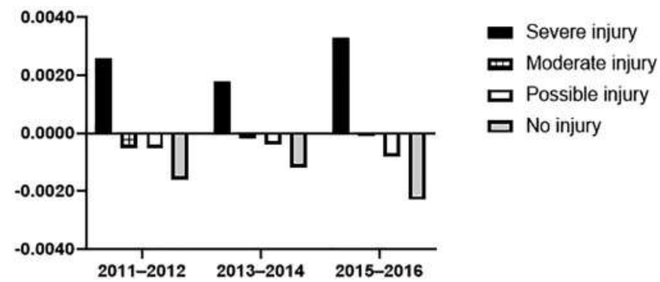


Fig. 9. Marginal effects of the indicator variable of exceeding speed limit.

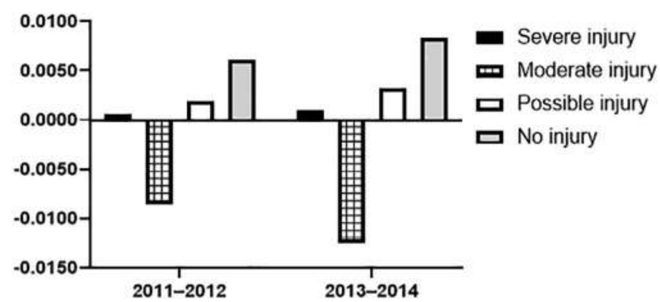


Fig. 10. Marginal effects of the indicator variable of afternoon time of day.

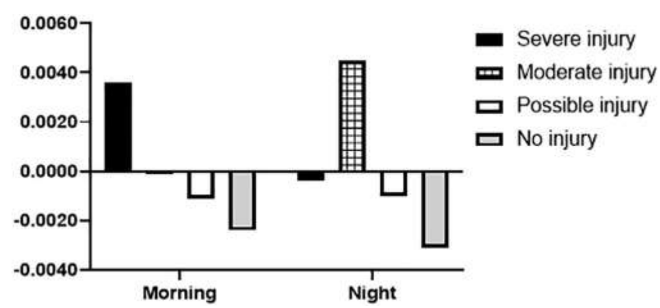


Fig. 11. Marginal effects of the indicator variables of morning and nighttime windows.

that fatigue and driving under influence at such time window are unlikely to be contributing factors.

6.5. Weather and lighting attributes

Regarding lighting conditions, the indicator variable of dawn lighting condition showed an increase in the likelihood of sustaining severe injuries in 2013–2014 time period model. Presumably, this could be attributed to sunlight glare that exacerbates the visibility of drivers, particularly on RTU horizontal curves.

Specific weather conditions were also found to have an effect on injury severity of crashes. Specifically, two indicators of weather conditions: clear and snowy weather were found to produce significant impacts. In 2011–2012 time period model, the increase in the probability of having severe injuries was conditioning that crashes occurred in clear weather. Driving at higher speed in such weather conditions could be a possible reason. In contrast, snowy conditions were found to reduce the probability of severe injuries in 2013–2014 time period model. Again, this finding may reflect driver behavior under such conditions, in which drivers tend to be tentative and driving carefully to compensate for these weather conditions. Fig. 12 clearly shows the difference in the effects of clear and snowy weather conditions on injury severity sustained.

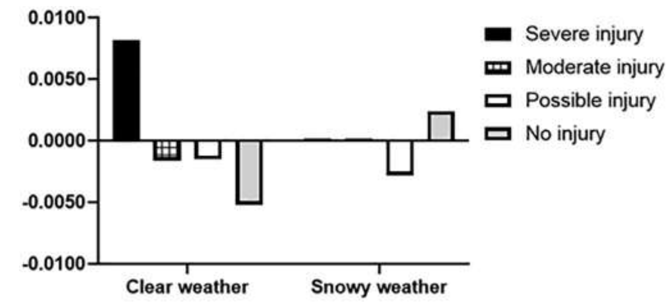


Fig. 12. Marginal effects of the indicator variables of weather conditions.

6.6. Heterogeneity in means and variances of random parameters

In the 2011–2012 time period model, belted drivers and crashes occurring on weekdays turned out to be statistically significant and produced a random parameter with significant heterogeneity in means (as seen in Table 4). The means of these parameters increased if a head-on collision was involved. These findings indicate that head-on collisions are extremely dangerous and pronounced on RTU horizontal curves. That being said, effective treatments on such locations are required to curb such crashes. For belted drivers, there was a significant increase in the mean of this variable if crashes involved a single vehicle, meaning that moderate injuries are highly likely to be suffered compared to other injury outcomes if crashes involved a single vehicle.

For crashes occurring on weekdays, the means of this variable increased if older drivers (over 65 years) were involved and crashes occurring on dry roadway surface conditions. This will result in an increase in severe and moderate injuries for crashes involving older drivers (over 65 years) and crashes occurring on dry roadway surface conditions, respectively (see Table 4). These findings could be interpreted as direct implications of physiological deterioration due to aging, which reduces reaction times, visibility, and driving skills. Also, drivers' overconfidence in dry surface conditions could be am underlying reason for moderate injuries from crashes occurring on dry roadways compared to that on wet conditions.

In terms of the heterogeneity in the variances, the variances of the two random parameters (belted drivers and crashes occurring on weekdays) increased if crashes involved female drivers and deployed airbags.

Table 5 clearly shows that in 2013–2014 two factors were found to produce a random parameter with heterogeneity in means; belted drivers and losing control of vehicle. For belted drivers and losing control of vehicle, drunk drivers resulted in an increase in their means, making severe injuries more likely to be incurred by drunk drivers in crashes along RTU horizontal curves compared to sober drivers. The lack of consciousness and deterioration in driving abilities could be reasons.

Table 5 also shows that the mean of belted drivers increased if female drivers were involved, which resulted in an increase in the possibility of sustaining possible injuries among female drivers relative to male drivers. Again, differences in risk taking behavior between males and females, meaning that females tend to be less risky compared with males, could substantiate such findings. As regards to belted driver, the mean of this variable decreased if crashes occurred in the winter season (between January and April). This finding proves that the probability of severe injury resulted from crashes occurring in winter along RTU horizontal curves is reduced compared to crashes in other seasons. Weather conditions associated with winter, such as heavy rain, snow, and fog, require paying more attention and being alert. As a result, crashes occurring in the winter along RTU horizontal curves tend to be less severe compared to crashes in other seasons.

As for the variance of belted drivers, the variance of crashes occurring on roadways with 45 mph speed limit is increased.

In the 2015–2016 time period model, four variables were found to be random and produced heterogeneity in means: wet roadway surface condition, belted drivers, young drivers, and middle-aged drivers. Table 6 shows that the means of belted drivers and young drivers increased if a single vehicle was involved, meaning that severe injuries are more likely being sustained in crashes involved single vehicle along RTU horizontal curves relative to crashes involved two or more vehicles. In contrast, the indicator variable of single vehicle decreased the means of middle-aged drivers and crashes occurred on wet roadway surfaces, making severe injuries less likely to be suffered compared to crashes involved two or more vehicles. The variation in injury severity sustained among drivers (young vs. middle-aged) in crashes involved single vehicle could reflect the role of age and driving experience in navigating curves and reducing injuries if crash occurred.

For the variances of belted drivers, young drivers, and losing control of vehicle, the variance of indicator of the deployed airbag is increased. Similarly, for belted drivers, the variance of the indicator variable of overturn crashes is increased.

7. Summary and conclusions

Among hazardous locations that need extensive research efforts is the safety of RTU horizontal curves. As such, the main interest of the current study is to shed some light on the factors that influence injury severity of crashes at such locations. Also, it is widely recognized that driver behavior and other attributes such as roadway geometry, driving skills, the way of being distracted are changing over time. As a result, accounting for temporal stability of these factors would be extremely important to correctly infer from the obtained estimates. Addressing temporal stability is not enough if the shortcomings in the traditional crash data gathered by police officers are not considered. Considering that, a mixed logit model with heterogeneity in means and variances of the random parameters was utilized to capture contributing factors to crashes occurred along RTU horizontal curves. To achieve the goal of this paper, crash data of Oregon for crashes occurred between 2011 and 2016 was used. Then, crash data was split into three time periods: 2011–2012, 2013–2014, and 2015–2016 to explore temporal stability of the contributing factors.

The estimation results reveal the hypothesis that the determinants of injury severity of crashes along RTU horizontal curves are stable over time must be rejected. This has been conducted by using a series of likelihood ratio tests. Older drivers with more than 65 years found to be significant and relatively stable in three time periods. However, young drivers with 30 years and younger and middle-aged drivers are temporally instable. This could be in part due to the rapid change in technology (such as cellphone apps) and how it impacts young and middle-aged drivers over time by means of distraction compared to older drivers who are away from this advancing technology. The estimation results also show that some variables have common effects in all or two of time period models, however, the marginal effects of these factors are significantly different. For example, distracted drivers were found to increase moderate injuries in 2015–2016 time period model while decreasing moderate injuries in 2013–2014 time period model. Continuous change in drivers' behavior such as distraction engagement, driving skills, risk taking behavior, responding to a stimulus, perception-reaction time, advanced technology equipped with new vehicles are potential reasons underlying temporal instability.

The findings can help and insight traffic authorities and official to improve safety of horizontal curves, particularly those with two-lanes, two-way undivided at rural roadways. For instance, based on the findings of this research, head-on collisions are more pronounced on RTU horizontal curves. Accordingly, geometric design of hazardous curves can be remediated by installing centerline rumble strips and considering speed limits at curves with high crash frequency. Regarding driver behavior factors, driving under the influence of alcohol and exceeding speed limit should be targeted with strict law enforcement. One approach of improvement is deploying curve signs effectively to reduce crashes and resulted injuries. Also, the findings reveal that an obvious temporal shift in crash risks, particularly behavioral factors is existing due to the implications of evolving technologies equipped with vehicles that could potentially increase the chances of distracted driving, the advancing in communication devices (cell phones), and risk-taking behaviors. Accordingly, addressing the temporal stability of risk factors of crashes along RTU horizontal curves can highly improve the safety of these hazardous locations.

It should be noted that some important factors are lacked in the used crash data such as curvature of horizontal curves, shoulder width and type, and pavement characteristics. Therefore, one could recommend for future studies to collect comprehensive crash data to gain further insights.

Acknowledgements

The support provided by the Oregon Department of Transportation (ODOT) by providing crash data is highly and gratefully acknowledged.

References

Al-Bdairi, N. S. S. (2020). Does time of day matter at highway work zone crashes? *Journal of Safety Research, 73*, 47–56.
Al-Bdairi, N. S. S., Behnood, A., & Hernandez, S. (2020). Temporal stability of driver injury severities in animal-vehicle collisions: A random parameters with heterogeneity in means (and variances) approach. *Analytic Methods in Accident Research, 26* 100120.
Al-Bdairi, N. S. S., & Hernandez, S. (2017). An empirical analysis of run-off-road injury severity crashes involving large trucks. *Accident Analysis & Prevention, 102*, 93–100.
Al-Bdairi, N. S. S., Hernandez, S., & Anderson, J. (2018). Contributing factors to run-off-road crashes involving large trucks under lighted and dark conditions. *Journal of Transportation Engineering, Part A: Systems, 144*, 1–9.
Alnawmasi, N., & Mannering, F. (2019). A statistical assessment of temporal instability in the factors determining motorcyclist injury severities. *Analytic Methods in Accident Research, 22*, 1–20.
Anarkooli, A. J., & Hosseinlou, M. H. (2016). Analysis of the injury severity of crashes by considering different lighting conditions on two-lane rural roads. *Journal of Safety Research, 56*, 57–65.
Anarkooli, A. J., Persaud, B., Hosseinpour, M., & Saleem, T. (2019). Comparison of univariate and two-stage approaches for estimating crash frequency by severity—Case study for horizontal curves on two-lane rural roads. *Accident Analysis and Prevention, 129*, 382–389.
Banihashemi, M. (2016). Effect of horizontal curves on urban arterial crashes. *Accident Analysis and Prevention, 95*, 20–26.
Behnood, A., & Al-Bdairi, N. S. S. (2020). Determinant of injury severities in large truck crashes: A weekly instability analysis. *Safety Science, 131*, 1–10.
Behnood, A., & Mannering, F. (2019). Time-of-day variations and temporal instability of factors affecting injury severities in large-truck crashes. *Analytic Methods in Accident Research, 23*, 100102.
Behnood, A., & Mannering, F. (2017a). Determinants of bicyclist injury severities in bicycle-vehicle crashes: A random parameters approach with heterogeneity in means and variances. *Analytic Methods in Accident Research, 16*, 35–47.
Behnood, A., & Mannering, F. (2017b). The effect of passengers on driver-injury severities in single-vehicle crashes: A random parameters heterogeneity-in-means approach. *Analytic Methods in Accident Research, 14*, 41–53.
Behnood, A., & Mannering, F. L. (2017). The effects of drug and alcohol consumption on driver injury severities in single-vehicle crashes. *Traffic Injury Prevention, 18* (5), 456–462.
Behnood, A., & Mannering, F. L. (2016). An empirical assessment of the effects of economic recessions on pedestrian-injury crashes using mixed and latent-class models. *Analytic Methods in Accident Research, 12*, 1–17.
Behnood, A., & Mannering, F. L. (2015). The temporal stability of factors affecting driver-injury severities in single-vehicle crashes: Some empirical evidence. *Analytic Methods in Accident Research, 8*, 7–32.
Behnood, A., Roshandeh, A. M., & Mannering, F. L. (2014). Latent class analysis of the effects of age, gender, and alcohol consumption on driver-injury severities. *Analytic Methods in Accident Research, 3-4*, 56–91.

Calvi, A. (2015). A study on driving performance along horizontal curves of rural roads. *Journal of Transportation Safety & Security, 7*.
Choudhari, T., & Maji, A. (2019). Socio-demographic and experience factors affecting drivers' runoff risk along horizontal curves of two-lane rural highway. *Journal of Safety Research, 71*, 1–11.
Dhahir, B., & Hassan, Y. (2019a). Using horizontal curve speed reduction extracted from the naturalistic driving study to predict curve collision frequency. *Accident Analysis and Prevention, 123*, 190–199.
Dhahir, B., & Hassan, Y. (2019b). Probabilistic, safety-explicit design of horizontal curves on two-lane rural highways based on reliability analysis of naturalistic driving data. *Accident Analysis and Prevention, 123*, 200–210.
Farah, H., Daamen, W., & Hoogendoorn, S. (2019). How do drivers negotiate horizontal ramp curves in system interchanges in the Netherlands? *Safety Science, 119*, 58–69.
Fitzsimmons, E. J., Lindheimer, T. E., Schrock, S. D., & Gonterwitz, K. (2012). Evaluation of large truck crashes at horizontal curves on two-lane rural highways in Kansas. *Presented at 92nd annual meeting of the transportation research board*.
Gabauer, D. J., & Li, X. (2015). Influence of horizontally curved roadway section characteristics on motorcycle-to-barrier crash frequency. *Accident Analysis and Prevention, 77*, 105–112.
Geedipally, S. R., Pratt, M. P., & Lord, D. (2019). Effects of geometry and pavement friction on horizontal curve crash frequency. *Journal of Transportation Safety & Security, 11*, 167–188.
Gooch, J. P., Gayah, V. V., & Donnell, E. T. (2018). Safety performance functions for horizontal curves and tangents on two lane, two way rural roads. *Accident Analysis and Prevention, 120*, 28–37.
Gooch, J. P., Gayah, V. V., & Donnell, E. T. (2016). Quantifying the safety effects of horizontal curves on two-way, two-lane rural roads. *Accident Analysis and Prevention, 92*, 71–81.
Jamson, S., Benetou, D., & Tate, F. (2015). The impact of arc visibility on curve negotiation. *Advances in Transportation Studies, 37*, 79–92.
Jesna, N. M., & Anjaneyulu, M. V. L. R. (2016). Reliability analysis of horizontal curves on two lane highways. *Transportation Research Procedia, 17*, 107–115.
Kim, J.-K., Ulfarsson, G. F., Kim, S., & Shankar, V. N. (2013). Driver-injury severity in single-vehicle crashes in California: A mixed logit analysis of heterogeneity due to age and gender. *Accident Analysis and Prevention, 50*, 1073–1081.
Kockelman, K. M., & Kweon, Y.-J. (2002). Driver injury severity: An application of ordered probit models. *Accident Analysis and Prevention, 34*, 313–321.
Lee, C., & Li, X. (2014). Analysis of injury severity of drivers involved in single- and two-vehicle crashes on highways in Ontario. *Accident Analysis and Prevention, 71*, 286–295.
Lord, D., Brewer, M. A., Fitzpatrick, K., Geedipally, S. R., Peng, Y. (2011). Analysis of roadway departure crashes on two-lane rural roads in Texas. Austin, Texas.
Mannering, F. (2018). Temporal instability and the analysis of highway accident data. *Analytic Methods in Accident Research, 17*, 1–13.
Mannering, F. L., Shankar, V., & Bhat, C. R. (2016). Unobserved heterogeneity and the statistical analysis of highway accident data. *Analytic Methods in Accident Research, 11*, 1–16.
McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics, 15*, 447–470.
Milton, J. C., Shankar, V. N., & Mannering, F. L. (2008). Highway accident severities and the mixed logit model: An exploratory empirical analysis. *Accident Analysis & Prevention, 40*, 260–266.
National Highway Traffic Safety Administration (NHTSA). (2019). Traffic safety facts 2017: A compilation of motor vehicle crash data (No. DOT HS 812 806). Washington DC.
Oregon Department of Transportation (ODOT). (2019). 2017 Oregon State Highway Crash Rate Tables. Salem, OR.
Pahukula, J., Hernandez, S., & Unnikrishnan, A. (2015). A time of day analysis of crashes involving large trucks in urban areas. *Accident Analysis and Prevention, 75*, 155–163.
Pratt, M. P., Geedipally, S. R., Wilson, B., Das, S., Brewer, M., & Lord, D. (2018). Pavement safety-based guidelines for horizontal curve safety (No. FHWA/TX-18/0-6932-R1). College Station, Texas.
Rakotonirainy, A., Chen, S., Scott-Parker, B., Loke, S. W., & Krishnaswamy, S. (2015). A novel approach to assessing road-curve crash severity. *Journal of Transportation Safety and Security, 7*, 358–375.
Russo, B. J., Savolainen, P. T., Schneider, W. H., IV, & Anastasopoulos, P. C. (2014). Comparison of factors affecting injury severity in angle collisions by fault status using a random parameters bivariate ordered probit model. *Analytic Methods in Accident Research, 2*, 21–29.
Schneider, W. H., Savolainen, P. T., & Zimmerman, K. (2009). Driver injury severity resulting from single-vehicle crashes along horizontal curves on rural two-lane highways. *Transportation Research Record, 2102*, 85–92.
Seraneeprakarn, P., Huang, S., Shankar, V., Mannering, F., Venkataraman, N., & Milton, J. (2017). Occupant injury severities in hybrid-vehicle involved crashes: A random parameters approach with heterogeneity in means and variances. *Analytic Methods in Accident Research, 15*, 41–55.
Torbic, D. J., Harwood, D. W., Gilmore, D. K., Pfefer, R., Neuman, T. R., Slack, K. L., & Hardy, K. K. (2004). NCHRP REPORT 500: Guidance for implementation of the AASHTO strategic highway safety plan. Volume 7: A Guide for Reducing Collisions on Horizontal Curves. Washington, D.C.
Wang, B., Hallmark, S., Savolainen, P., & Dong, J. (2017). Crashes and near-crashes on horizontal curves along rural two-lane highways: Analysis of naturalistic driving data. *Journal of Safety Research, 63*, 163–169.
Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2011). *Statistical and econometric methods for transportation data analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.
Wu, J., & Xu, H. (2017). Driver behavior analysis for right-turn drivers at signalized intersections using SHRP 2 naturalistic driving study data. *Journal of Safety Research, 63*, 177–185.
Wu, L., Dadashova, B., Geedipally, S., Pratt, M. P., & Shirazi, M. (2019). Using naturalistic driving study data to explore the association between horizontal curve safety and operation on rural two-lane highways. *Journal of Transportation Safety and Security*, 1–18.
Xin, C., Wang, Z., Lee, C., Lin, P.-S., Chen, T., Guo, R., & Lu, Q. (2019). Development of crash modification factors of horizontal curve design features for single-motorcycle crashes on rural two-lane highways: A matched case-control study. *Accident Analysis and Prevention, 123*, 51–59.
Xin, C., Wang, Z., Lin, P.-S., Lee, C., & Guo, R. (2017). Safety effects of horizontal curve design on motorcycle crash frequency on rural, two-lane, undivided highways in Florida. *Presented at 96th annual meeting of the transportation research board*.
Ye, F., & Lord, D. (2014). Comparing three commonly used crash severity models on sample size requirements: Multinomial logit, ordered probit and mixed logit models. *Analytic Methods in Accident Research, 1*, 72–85.
Zou, W., Wang, X., & Zhang, D. (2017). Truck crash severity in New York City: An investigation of the spatial and the time of day effects. *Accident Analysis and Prevention, 99*, 249–261.

Nabeel Saleem Saad Al-Bdairi is a lecturer at the University of Wasit, Kut, Iraq where he teaches undergraduate courses in civil engineering and conducts transportation safety research. He earned his Ph.D. from Oregon State University with a focus on large truck safety. His research interests encompass a broad domain of transportation safety, network modeling and data analysis (big data, spatial analyses, heterogeneity, etc.). His primary emphasis is on transportation safety analysis by using advanced econometric methods.

Ali Behnood, Ph.D. is currently a Research Associate at Purdue University where he also received his PhD in Civil Engineering. As a multidisciplinary researcher, he works on various topics related to Civil Engineering and Transportation Engineering. One of the main areas of his research includes the analysis of traffic safety using a wide range of econometric methods and artificial intelligence-based techniques.

Journal of Safety Research 76 (2021) 56–63

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Children's fear in traffic and its association with pedestrian decisions

Huarong Wang [a,*,1], Casie Morgan [b], Dongqian Li [a], Rong Huang [a], David C. Schwebel [b]

[a] Department of Traffic Psychology, Institute of Special Environmental Medicine, Nantong University, 9 Seyuan Road, Nantong, Jiangsu Province 226019, China
[b] Department of Psychology, University of Alabama at Birmingham, 1300 University, Blvd, CH 415, Birmingham, AL 35294, USA

ARTICLE INFO

Article history:
Received 9 April 2020
Received in revised form 16 July 2020
Accepted 23 November 2020
Available online 15 December 2020

Keywords:
Pedestrian
Children
Fear
Traffic situation
Pedestrian decision

ABSTRACT

Introduction: Research on risk for child pedestrian injury risk focuses primarily on cognitive risk factors, but emotional states such as fear may also be relevant to injury risk. The current study examined children's perception of fear in various traffic situations and the relationship between fear perception and pedestrian decisions. *Method:* 150 children aged 6–12-years old made pedestrian decisions using a table-top road model. Their perceived fear in the pedestrian context was assessed. *Results:* Children reported greater emotional fear when they faced quicker traffic, shorter distances from approaching traffic, and red rather than green traffic signals. Children who were more fearful made safer pedestrian decisions in more challenging traffic situations. However, when the least risky traffic situation was presented, fear was associated with more errors in children's pedestrian decisions: fearful children failed to cross the street when they could have done so safely. Perception of fear did not vary by child age, although safe pedestrian decisions were more common among the older children. *Conclusions:* Children's emotional fear may predict risk-taking in traffic. When traffic situations are challenging to cross within, fear may appropriately create safer decisions. However, when the traffic situation is less risky, feelings of fear could lead to excessive caution and inefficiency. *Practical applications:* Child pedestrian safety interventions may benefit by incorporating activities that introduce realistic fear of traffic risks into broader safety lessons.

1. Introduction

Worldwide, road traffic injury (RTI) is the leading cause of death for children over five years old (WHO, 2018), and pedestrian injury is the most common type of child RTI mortality. According to Global Burden of Disease data estimates, 34,000 children ages 5–14 died from road traffic crashes in 2017; 25,000 (74%) of those deaths were child pedestrians. In China, the focus of this study, 68% of children ages 5–14 who die in RTI are pedestrians (Institute for Health Metrics and Evaluation, 2019).

Given these data, behavioral scientists have examined factors that may increase risk of children's pedestrian injuries (Ampofo-Boateng & Thomson, 1991; Kovesdi & Barton, 2013; Meir, Oron-Gilad, & Parmet, 2015; Meir, Parmet, & Oron-Gilad, 2013; Schwebel, Davis, & O'Neal, 2012; Schwebel & Gaines, 2007; Wazana, Krueger, Raina, & Chambers, 1997). One prominent risk factor is children's cognitive-perceptual skills. Since children's cognitive-perceptual skills naturally develop through the early and middle childhood years, they demonstrate inadequate knowledge about road safety rules and practices (Dong et al., 2011; Koekemoer, Gesselleen, Niekerk, Govender, & As, 2017), misperceive risk from oncoming traffic (Demetre, 1997; Poudel-Tandukar, Nakahara, Ichikawa, Poudel, & Jimba, 2007), and demonstrate inadequate attention and visual searching abilities amidst traffic (Tabibi & Pfeffer, 2007; Whitebread & Neilson, 2000). Deficits in these cognitive-perceptual skills increase risk for traffic-related injuries.

Scholars in the broader child injury prevention literature have recently suggested, however, that emotional state is relevant to child injury risk as well as cognitive skills (Morrongiello, Corbett, Switzer, & Hall, 2015; Morrongiello & Lasenby-Lessard, 2006; Morrongiello & Matheis, 2007). In pedestrian settings, the role of fearful emotions in influencing children's decision-making or behavior in traffic is poorly understood. Basic social learning principles suggest children who experience, observe, or learn about others being injured or in danger while crossing the street, might develop fearful emotions surrounding street-crossing, leading the children to avoid or take caution when engaging in street-crossing behavior (Bandura, 1977; Skinner, 1969). Results from the broad child injury literature support this hypothesis

* Corresponding author at: No.9 Seyuan Road, Chongchuan District, Nantong city, Nantong University, 226019, China.
E-mail address: yeluo801004@163.com (H. Wang).
[1] Huarong Wang and Casie Morgan contributed equally to this work.

https://doi.org/10.1016/j.jsr.2020.11.010
0022-4375/

(Morrongiello & Matheis, 2004), as do results in the driving literature (Bouffard, 2014; Lu, Xie, & Zhang, 2013).

Other evidence from traffic safety contradicts such patterns, however, suggesting fearful individuals – usually conceptualized from a temperament or personality perspective rather than an emotional state of fear – may take greater risks in traffic situations. Shen, McClure, and Schwebel (2015) found, for example, significant relations between temperamental fear and children's risky pedestrian behavior in a virtual reality traffic environment. In that study, fearful children were more likely to hesitate prior to crossing, and thus entered and crossed within smaller traffic gaps and had increased risk of virtual collisions. Similarly, some adult driving research supports fear as an influence on greater risk-taking in traffic. Taylor, Deane, and Podd (2007) reported, for example, that fearful drivers made significantly more mistakes during an on-road driving assessment than non-fearful ones.

The present study used a table-top road model to manipulate traffic speeds, distances, and traffic signal lights and examine children's perception of fear in various traffic situations and across different age groups. We hypothesized children would report different degrees of fear in various traffic scenarios, with greater perception of fear under more challenging or threatening traffic scenarios (based on vehicle speed and distance, and red instead of green traffic lights). We also considered whether fear ratings would influence decisions to cross. As a secondary question, we examined developmental influences, anticipating older children would make safer decisions and experience less fear in pedestrian settings.

2. Methods

2.1. Participants

Participants were recruited from an elementary school in Nantong City, China and included 150 children (75 females, 75 males) aged 6–12-years-old ($M = 9.14$ years, $SD = 1.89$; range = 6.17–12.92 years). Participants were recruited from the 1st grade ($N = 50$; 25 girls and 25 boys; $M = 7.39$ years, $SD = 0.53$, range = 6.17 years to 8.25 years), 3rd grade ($N = 50$; 25 girls and 25 boys; $M = 9.44$ years, $SD = 0.87$, range = 8.42 years to 11.25 years), and 5th grade ($N = 50$; 25 girls and 25 boys; $M = 11.31$ years, $SD = 0.65$, range = 9.42 years to 12.92 years). Thirty-seven children (25%) reported walking as their most common means of transportation to school, and 2 students (1%) reported taking public transportation. The remaining children (111 students; 74%) took a family vehicle/motorcycle to school most often. The following exclusion criteria were applied: (a) history of self-reported pedestrian traffic-related injury; (b) impaired vision or hearing, as reported by the classroom teacher; and (c) attentional or other cognitive impairment, as reported by the classroom teacher. No children were excluded for these reasons. The study protocol was reviewed and approved by the Nantong University Academic Ethics Committee and the Primary Education Office of participating schools. Written parental consent was obtained prior to each child's participation.

2.2. Materials

2.2.1. Table-top road model

Previous studies demonstrate table-top road models to be both feasible and valid for studying children's road safety (e.g., Albert & Dolgin, 2010; Ampofo-Boateng & Thomson, 1991; Barton & Schwebel, 2007; Thomson, 1997; Twisk, Vlakveld, Mesken, Shope, & Kok, 2013). Building from these previous studies, we designed a table-top model of an urban Chinese community setting (See Fig. 1). The model, which measured 120 cm × 100 cm, included an apartment complex, school, hospital, park, bank, gas station, auto repair shop, supermarket, hotel, playground, and mall. The road traffic environment was simulated by multiple 8 cm-long toy cars and by 7 cm-high traffic lights which could be manually switched between red and green lights. Two felt figures of children (4.5 cm tall; one boy and one girl) were used, with gender-matching to participants. The table-top model was affixed to a table about 1 meter high.

2.2.2. Pedestrian task protocol

Each participant engaged in traffic situations that manipulated three variables: traffic light signal (red vs. green), vehicle distance (far vs. near), and vehicle speed (fast vs. slow). Each of these situations was considered to be a singular variable for data analysis and each variable had two levels. Based on the size of the model, 48 cm and 13 cm were set as the "far vehicle distance" and "near vehicle distance," respectively, and 1.4 m/s and 0.5 m/s were set as the "fast vehicle speed" and "slow vehicle speed," respectively. Vehicle distance and speed were manipulated within the same trials (that is, children experienced cars traveling far and fast, far and slow, near and fast, and near and slow) while traffic light signals were manipulated in separate trials.

In total, therefore, children completed six types of road-crossing scenarios: (a) red light signal, (b) green light signal, (c) slow car approaching from near distance, (d) slow car approaching from far distance, (e) fast car approaching from near distance, and (f) fast car approaching from far distance. The two traffic light signal situations were presented four times each and the four vehicle approaching situations were presented two times each. Thus, each child engaged in 16 total crossings.

Automated "pull-back" toy cars, which had their speed controlled through an internal wind-up spring that created consistent speeds across trials and across children, were used to represent moving traffic. They were placed by research assistants at the required distance for each situation. Two identical-looking cars were used, one for "fast" trials and the other for "slow" trials; the internal spring mechanism varied, creating the different speeds. Research assistants manually manipulated the speed of the car by placing the appropriate car on the tabletop, and children judged safety by watching and deciding if they would cross the street given the location and speed of the model car. During all scenarios, other cars were placed in stationary locations, representing travel on different roads. They simulated typical traffic scenarios and discouraged children from selecting alternate routes with less traffic to reach their destination.

To increase realism and diversify the experimental experience, the 16 road-crossing situations were randomly matched across participants to 8 different pedestrian travel scenarios, each of which was presented twice for each participant: (1) Supermarket → Bus Station; (2) Bus Station → Supermarket; (3) Shopping Mall → Park; (4) Park → Shopping Mall; (5) Telephone Booth → Playground; (6) Playground → Telephone Booth; (7) Park → Gas Station; and (8) Gas Station → Park. Each task was selected to meet the following criteria: (a) there was more than one route available to travel from the starting point to the destination, and (b) only one of the available routes was the shortest, defined as children needed to cross *only* two roads to reach the destination. Randomization was accomplished through a system of random-number draws.

Children were introduced to each travel scenario in the context of a short vignette story. In each story, the boy/girl had the task of meeting his/her friend. For example, for the Shopping Mall → Park task, children heard: *Today is* [date and time of the experiment]. *This child* [researcher displays gender-matched toy doll] *is going to the park to meet his/her* [gender-matched] *friend. His/her mom is going shopping in the mall, and while she is shopping, she lets the*

Fig. 1. Table-top road model.

boy/girl go to the park by him/herself. Would you please help this boy/-girl get to the park safely? While each story was being told, a second researcher placed the experimental props (the traffic lights or vehicles) in appropriate locations on the model, reflecting the scenario presented and the roads children would likely choose to reach the destination.

2.3. Procedure

Each child was brought by his/her classroom teacher to a vacant classroom at the school, where two trained research assistants guided the child through the experiment. One researcher arranged the vehicles and traffic lights on the table-top model corresponding to the randomization scheme. The second researcher read each pedestrian vignette and then asked the child to help the doll reach the destination safely by moving the doll across the roads.

After each child listened to the pedestrian vignette, she/he was asked to report her/his route plan and be ready to cross the road. Just before the child began moving the doll to cross the road, the researcher set one of the vehicles in motion toward the doll and the other research assistant asked the child two questions: (1) would you decide to cross the road at this moment? and (2) if you crossed the road at this moment, how scared would you be? After answering the questions, researchers stopped the moving vehicle and allowed the child to move the doll across the road. For approaching vehicle tasks, therefore, children's decisions were made as the car moved on the model street toward the doll. For the traffic light tasks, children were asked to answer the two questions just before moving the doll to cross the road, and no vehicles moved.

To reduce the influence on our outcome measures of children randomly selecting safe or unsafe crossing options given the binary choices presented, following each crossing the researcher asked the child to explain why he or she chose to cross (see Ampofo-Boateng & Thomson, 1991; Morrongiello & Matheis, 2004; Whitebread & Neilson, 2000). Children's responses were coded as either appropriate (1 point; e.g., *I didn't cross because the car was moving fast toward me*) or illogical (0 points; e.g., *I didn't cross because my legs were tired*). In rare instances when the researcher was unsure how to score a child's response (<3%), the response was recorded in writing and then discussed with the research team to reach a consensus score.

Fear responses ("How scared would you be?") were recorded on a 5-point scale (from 0 to 4). To ensure children understood the response guidelines, at the start of the session the researcher offered the following example: *Let's imagine you are rating how cold you feel. In summer, it's very hot and you will not feel cold. You should answer 0 for how cold you feel. In autumn, it's cooler, and you might feel a little chilly. So in autumn, you can rate your coldness with 1 point. In winter, the temperature is much lower and you will feel cold. So you can rate winter as 2 points. On some days in winter there is thick snow. On those days, you will feel very cold. You might rate those cold and snowy days with 3 points. And what if you fall into an icy river during winter? Then you will feel extremely cold. You would rate that as 4 points.* Following that example, children were instructed to use similar methods to assess their level of fear in each scenario. Not feeling fear was rated as 0; a little fear was rated as 1; more fear was rated as 2; very fearful was rated as 3; and extremely fearful was rated as 4. The study took about 45–60 minutes for each participant, and each child was compensated with a selection of small gifts, such as school supplies.

2.4. Outcome measures

Three outcome measures were obtained from each pedestrian story: (a) the child's decision to cross at an appropriate time (details in next paragraph), (b) the appropriateness of the child's reason for crossing or not crossing, and (c) how fearful the child felt to cross. The decision to cross was coded as safe (1) or unsafe (0). The reason to cross was coded as either logical and appropriate (1) or illogical and inappropriate (0). We merged these two scores (decision to cross and appropriateness of reason) into a single composite variable of pedestrian decisions by summing the two scores and yielding a crossing safety outcome score ranging from 0 to 2 for each crossing trial. For example, the child scored the minimum of zero if they were rated with a 0 for both their decision and reason scores. The child scored the maximum of two if they demonstrated a safe decision to cross with a logical reason for crossing.

Decisions about whether the child chose to cross at an appropriate time were based on the arrival time of the approaching toy car (about 0.09 s, 0.26 s, 0.34 s and 0.96 s, respectively, for the four combinations of approaching vehicle speed and distance) and the length of time needed for the child to cross the road (about 0.12 s, according to children's average walking pace in this study, 1.3 m/s, and the width of simulated road). Decisions were considered appropriate if children chose to cross when the vehicle was moving slowly and was a far distance away, or if the children chose not to cross when the vehicle was in any of the three other conditions. Similarly, decisions were considered appropriate when children crossed when they saw a green traffic light or if they waited when they saw a red traffic light. Again, appropriate decisions were coded as 1 and inappropriate decisions were coded as 0. The fear rating was measured at each crossing and analyzed in its original form (range = 0–4).

Because children completed two trials at each vehicle speed/distance combination, the crossing safety and fear scores were each averaged across those trials to yield four "safety scores" for each participant, representing the four speed/distance combinations. Similarly, because children completed four trials at each traffic signal the safety and fear scores were averaged across the four trials for each of the two colored light signals.

2.5. Data analysis

We implemented a 2 (sex: female vs. male) * 3 (grade: 1st grade, 3rd grade vs. 5th grade) * 2 (vehicle distance: far vs. near) * 2 (vehicle speed: fast vs. slow) mixed factorial design and a 2 (sex: female vs. male) * 3 (grade: 1st grade, 3rd grade vs. 5th grade) * 2 (traffic lights: red vs. green) mixed factorial design separately. Grade and sex served as between-subject predictors, and vehicle speed, vehicle distance and traffic lights as within-subject predictors. Data cleaning excluded 2 children due to experimenter error during the trials. The remaining 148 children included 50 first-graders, 48 third-graders, and 50 fifth-graders. Preliminary analyses examined bivariate relations between sex and grade and the outcome measures; sex was omitted from subsequent analyses given null results in preliminary analyses. Primary analyses were conducted through a series of repeated measures ANOVA models with grade, vehicle distance, vehicle speed and traffic lights serving as the independent variables and the pedestrian decision and fear perception outcomes as the dependent variables. Both main effects and interactions between multiple factors were tested. Finally, Pearson correlations and a series of logistic regressions were conducted to analyze relations between children's fear and their pedestrian decisions.

3. Results

Our first objective was to evaluate children's fear perceptions across different traffic speeds and distances. Descriptive data appear in Table 1. A 3 (grade: 1st vs. 3rd vs. 5th) * 2 (vehicle distance: far vs. near) * 2 (vehicle speed: fast vs. slow) repeated measures ANOVA indicated statistically significant effects of vehicle distance (F_{distance} (1, 145) = 87.22, $p < 0.001$, $\eta^2_{p=}0.38$) and vehicle speed (F_{speed} (1, 145) = 340.69, $p < 0.001$, $\eta^2_{p=}0.70$) on children's fear perception. We also detected a statistically significant interaction effect between vehicle distance and speed, $F_{\text{distance}\times\text{speed}}$ (1, 145) = 10.32, $p < 0.01$, $\eta^2_{p=}0.07$. Simple effect tests showed children experienced more fear when the vehicle was near (M_{near} = 1.72, M_{far} = 2.24) and when vehicles were moving fast (M_{fast} = 2.69, M_{slow} = 1.27). There were no significant grade effects, F (2, 145) < 1.

Our second objective was to examine children's fear perception in traffic scenarios with red and green traffic lights. Descriptive data appear in Table 2. A 3 (grade: 1st vs. 3rd vs. 5th) * 2 (traffic lights: red vs. green) repeated measures ANOVA on fear perception was computed. A main effect of traffic lights was found, F(1, 145) = 413.53, $p < 0.001$, $\eta^2_{p=}0.74$, indicating children experienced more fear when they crossed the road with a red traffic light present than with a green traffic light present (M_{red} = 2.22, M_{green} = 0.09). There were no significant grade effects, F (2, 145) < 1, nor did interaction effects emerge between traffic lights and grade, F (2, 145) < 1.

Table 1
Means (Standard Deviation) of Fear Perception in Response to Different Traffic Scenarios and across Different Grade Groups.

Grade	Vehicle Distance: Far		Vehicle Distance: Near		Traffic Lights	
	Fast	Slow	Fast	Slow	Red	Green
First grade (n = 50)	2.70 (1.39)	0.97 (1.14)	2.93 (1.22)	1.65 (1.22)	2.25 (1.25)	0.16 (0.37)
Third grade (n = 48)	2.31 (1.14)	0.76 (0.76)	2.88 (1.16)	1.49 (0.95)	3.17 (1.31)	0.03 (0.11)
Fifth grade (n = 50)	2.49 (0.96)	1.07 (1.04)	2.80 (1.05)	1.67 (0.92)	2.24 (1.20)	0.09 (0.34)
Total (N = 148)	2.50 (1.18)	0.94 (0.99)	2.87 (1.14)	1.60 (1.04)	2.22 (1.25)	0.09 (0.30)

Table 2
Means (Standard Deviation) Score of Pedestrian Decisions across Different Traffic Scenarios and Grade Groups.

Participant Grade Groups	Approaching Vehicles Far		Approaching Vehicles Near		Traffic Lights	
	Fast	Slow	Fast	Slow	Red	Green
First grade (n = 50)	1.49 (0.53)	1.18 (0.60)	1.44 (0.55)	1.30 (0.64)	1.52 (0.47)	1.56 (0.45)
Third grade (n = 48)	1.66 (0.43)	1.42 (0.61)	1.75 (0.39)	1.51 (0.48)	1.67 (0.48)	1.80 (0.32)
Fifth grade (n = 50)	1.85 (0.31)	1.33 (0.55)	1.87 (0.30)	1.71 (0.46)	1.70 (0.47)	1.86 (0.28)
Total (N = 148)	1.66 (0.45)	1.31 (0.59)	1.68 (0.46)	1.51 (0.55)	1.63 (0.48)	1.74 (0.38)

A series of repeated measures ANOVAs yielded statistically significant effects of vehicle distance ($F_{distance}$ (1, 145) = 11.19, $p < 0.001$, $\eta^2_{p=}0.07$), vehicle speed (F_{speed} (1, 145) = 137.63, $p < 0.001$, $\eta^2_{p=}0.49$) and traffic lights ($F_{traffic\ lights}$ (1, 145) = 13.61, $p < 0.001$, $\eta^2_{p=}0.09$) on children's pedestrian decisions. A main effect of grade was also found, F (2, 145) = 9.89, $p < 0.001$, $\eta^2_{p=}0.12$). Bonferroni post-hoc analyses indicated children in 1st grade scored lower in safety than children in 3rd grade (M = 1.35 vs. 1.58) and in 5th grade (M = 1.35 vs. 1.69), and children scored higher in safety when the traffic light was green than when it was red (M = 1.74 vs. 1.63).

We also found statistically significant interaction effects between vehicle distance and speed, $F_{distance \times speed}$ (2, 145) = 6.64, $p < 0.05$, $\eta^2_{p=}0.04$. Simple effect tests showed that when the car was approaching quickly, children's crossing decision scores were similar for far and near distances (M = 1.66 vs. 1.68, $p > 0.05$). When the car was approaching slowly, however, children's crossing decision scores were lower when the approaching car was far rather than near (M = 1.31 vs. 1.51, $p < 0.01$). Overall, children's decision scores were safer when the approaching car was moving fast than when it was moving slow (M = 1.68 vs. 1.41, $p < 0.001$) (Fig. 2).

Our final objective was to consider relations between children's fear and their pedestrian decisions. Spearman correlations were computed first. As shown in Table 3, there were significant positive associations between fear perception and pedestrian decisions for the far and fast condition (r = 0.32, $p < 0.01$), the near and fast condition (r = 0.29, $p < 0.01$), and the near and slow condition (r = 0.51, $p < 0.01$), and a significant negative association between fear perception and pedestrian decisions in the far and slow condition (r = -0.59, $p < 0.01$). Significant associations also emerged between fear perception and pedestrian decisions during the red light condition (r = 0.17, $p < 0.05$; See Table 4) and the green light condition (r = -0.18, $p < 0.05$).

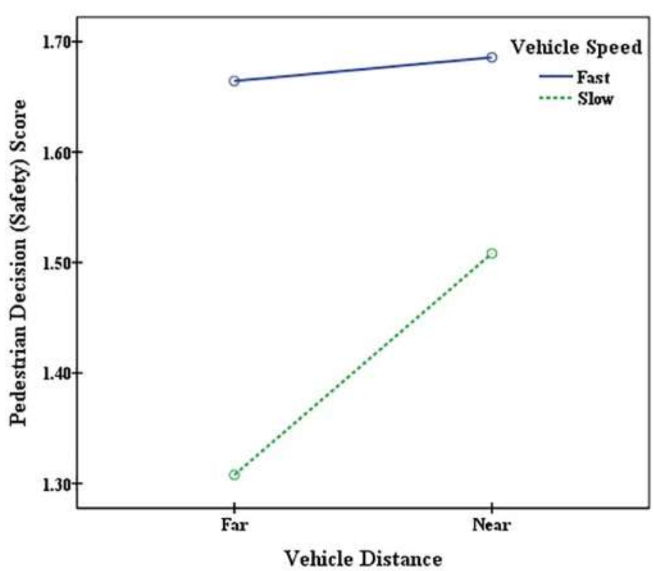


Fig. 2. Children's Pedestrian Decisions as a Function of Vehicle Distance and Vehicle Speed.

Last, a series of logistic regressions were computed with grade and fear perception as independent variables predicting each of the pedestrian decisions (see Tables 5 and 6). For this analysis, only pedestrian decisions (decision to cross rather than appropriateness of reason) was considered, and each decision was considered independently, without averaging across trials. Fear perception contributed significantly to pedestrian decisions in most models (far and slow vehicle condition, Wald χ^2 (df = 1, N = 149) = 15.11, $p < 0.01$; near and fast vehicle condition, Wald χ^2 (df = 1, N = 149) = 4.03, $p < 0.05$, near and slow vehicle condition, Wald χ^2 (df = 1, N = 149) = 18.27, $p < 0.01$, and the red light traffic condition, Wald χ^2 (df = 1, N = 147) = 4.371, $p < 0.05$), and approached traditional significance levels for the far and fast pedestrian condition decisions, Wald (df = 1, N = 149) = 3.70, p = 0.054. Fear perception was not a significant predictor of pedestrian decisions in the green light traffic condition after controlling for grade.

4. Discussion

This study investigated relations between children's fear in various traffic scenarios and their pedestrian decisions. As hypothesized, children reported greater degrees of fear when faced with more challenging traffic scenarios (those with faster vehicle speeds and closer vehicle distances), as well as when they encountered red rather than green traffic signals. Results concerning relations between perceived fear and pedestrian decisions indicated that fearful children were more cautious in their behavior near traffic. In more challenging traffic situations, more fearful children made safer pedestrian decisions; their fear led to appropriate caution as vehicle approached. When the least risky traffic situation was presented, however, fear was associated with excessive caution and errors in children's pedestrian decisions; fearful children chose not to cross when they could have done so safely. We discuss each of these findings below.

As predicted, and consistent with previous work (Rosenbloom, Nemrodov, Ben-Eliyahu, & Eldror, 2008), our results showed that children reported more fear when they faced more dangerous traffic situations. This finding was true even among the youngest children we studied, indicating that by first grade, children may have developed sufficient cognitive skills to apprise traffic- related danger and risk (Ampofo-Boateng & Thomson, 1991; Meir et al., 2015, 2013).

Our results concerning the association between children's fear and their pedestrian decisions offer insights to help explain the apparent contradictions in the existing literature. Consistent with

Table 3
Correlations between Fear Perception and Pedestrian Decisions for Vehicle Speed/Distance.

	1	2	3	4	5	6	7
Fear: Vehicle moving fast, far distance	1						
Fear: Vehicle moving slow, far distance	0.41**	1					
Fear: Vehicle moving fast, near distance	0.69**	0.32**	1				
Fear: Vehicle moving slow, near distance	0.52**	0.58**	0.56**	1			
Decisions: Vehicle moving fast, far distance	0.32**	0.22**	0.19*	0.25**	1		
Decisions: vehicle moving slow, far distance	−0.30**	−0.59**	−0.19*	−0.27**	−0.28**	1	
Decisions: vehicle moving fast, near distance	0.21*	0.08	0.29**	0.14	0.40**	−0.09	1
Decisions: vehicle moving slow, near distance	0.30**	0.36**	0.35**	0.51**	0.48**	−0.53**	0.26**

Note: * p < 0.05; ** p < 0.01.

Table 4
Correlations between Fear Perception and Pedestrian Decisions in the Traffic Light Condition.

	Fear: Red Light	Fear: Green Light	Decisions: Red Light
Fear: Red Light	1		
Fear: Green Light	0.03	1	
Decisions: Red Light	0.17*	−0.10	1
Decisions: Green Light	0.09	−0.18*	0.60**

Note. * $p < 0.05$; ** $p < 0.01$.

findings from the broader child injury field, when faced with traffic situations that presented considerable risk, higher perception of fear yielded safer pedestrian decisions (Morrongiello & Matheis, 2004). This was true in three of the moving vehicle conditions (vehicles moving fast from a near distance, slow from a near distance, and fast from a far distance), as well as when children were presented the option of crossing the street when they faced a red (don't walk) signal. The result reflects a developmentally-appropriate fearful response to a challenging situation that threatens personal safety.

When children faced the comparatively safe traffic situation of a vehicle moving toward them at a far distance and slow speed, however, fear related to more errors in pedestrian decisions: fearful children failed to cross the road when they could have done so safely. We also found a non-significant trend for this behavior pattern when children faced crossing with a green traffic light, a result that was impacted mathematically by near-ceiling scores among the older age groups of children and small variances among all age groups. It seems likely that the most fearful children, whether driven by temperamental or situational fear, hesitate or pause out of fear and anxiety in situations when most children would cross without hesitation (Shen et al., 2015). These emotion-driven delays lead to cognitive indecision or contemplation, increasing risk to the child as the traffic continues to move toward the child, and leading to errors including inefficient road crossing and failure to cross the street when a safe crossing is possible. Such behavior may also increase traffic-related fear through cyclical patterns, as close calls will inevitably ensue and could increase fearful emotions.

Our results also offer new insights concerning the development of both pedestrian safety skills and traffic-related fear. As expected, older children made safer pedestrian decisions than younger ones (Ampofo-Boateng & Thomson, 1991; Barton, Ulrich, & Lyday, 2011). We did not find relations between child age and perceptions of fear in traffic, a finding that matches some previous findings (Derevensky, 1974; Ollendick, 1983) but not others (Rosenbloom et al., 2008).

The findings offer tantalizing opportunities for injury prevention efforts. Fear-appeal tactics are used frequently in other health promotion campaigns (Job, 1985, 1988) and in safety training programs (Peters, Ruiter, & Kok, 2012; Will, Sabo, & Porter, 2009). They yield varying results (Goldenbeld, Twisk, & Houwing, 2008; Ruiter, Abraham, & Kok, 2001) but, consistent with our findings, tend to be less effective when the threat-level (perceived susceptibility and severity of the threat) is low (Price et al., 2011). Our findings indicate perceived fear yielded greater safety for children in higher-risk pedestrian situations, a result that matches the fear-appeal intervention literature, which suggests such campaigns are most effective when they depict a significant and relevant threat to the target audience, and when they outline effective responses that appear feasible to accomplish (Ahmadi & Ytterstad, 2007; Lewis, Watson, & White, 2008).

Thus, if an intervention were designed to stimulate appropriate levels of fear of traffic-related injury among children, and if those children were given appropriate tools to manage that fear through self-efficacy to cross streets safely, we might yield increased safe pedestrian behaviors. Such efforts would need to be conducted cautiously, with the goals of increasing children's perceived susceptibility to pedestrian safety and the potential severity of such injuries, ultimately creating moderate levels of fear accompanied by lessons on how to cross streets safely and efforts to build self-efficacy to engage in street-crossing behaviors.

Although we conducted this study with scientific rigor, like all research our study had limitations. First, we simulated the traffic situations by using table-top road models to examine children's fear perception and pedestrian decisions. Simulations are used frequently in the field and have evidence of validity (Ampofo-Boateng & Thomson, 1991; Thomson, 1997), but it is unclear whether children's behavior in tabletop simulations match precisely what they would do in real-world situations. Second, we used self-report to

Table 5
Summary of Logistic Regression Analysis for Variables Predicting Safe and Not Safe Pedestrian Decisions with Vehicle Speed/Distance Manipulated.

	Far and Fast Condition			Far and Slow Condition			Near and Fast Condition			Near and Slow Condition		
Predictor	β	*SE* β	e^b (95% CI)	β	*SE* β	e^b (95% CI)	β	*SE* β	e^b (95% CI)	β	*SE* β	e^b (95% CI)
First Grade	−1.56**	0.46	0.21(0.09–0.51)	−0.68	0.49	0.51(0.19–1.32)	−1.89***	0.48	0.15(0.06–0.39)	−1.47**	0.47	0.23(0.09–0.58)
Third Grade	−0.93*	0.45	0.39(0.16–0.96)	0.28	0.45	1.32(0.54–3.21)	−0.82	0.49	0.44(0.17–1.15)	−0.81	0.44	0.44(0.19–1.06)
Fear Perception	0.30	0.16	1.35(1.00–1.83)	−1.07**	0.28	0.34(0.20–5.88)	0.34*	0.17	1.40(1.01–1.93)	0.84**	0.20	2.32(1.58–3.42)

Notes: Fifth grade is the reference category.
*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$

Table 6
Summary of Logistic Regression Analysis for Variables Predicting Safe and Not Safe Pedestrian Decisions with Traffic Signal Light Manipulated.

	Red Light Condition			Green Light Condition		
Predictor	β	*SE* β	e^b (95% CI)	β	*SE* β	e^b (95% CI)
First Grade	−1.12*	0.46	0.33 (0.13–0.80)	−0.1.52**	0.48	0.22 (0.09–0.56)
Third Grade	−0.33	0.48	0.72 (0.28–1.83)	−0.23	0.54	0.79 (0.28–2.29)
Fear Perception	0.31*	0.15	1.37 (1.02–1.85)	−0.73	0.61	0.48 (0.15–1.58)

Notes: Fifth grade is the reference category.
*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$

record children's fear perception in several traffic situations. It is unclear how accurate children's self-reported fear is (Gullone & Lane, 1997) and future research might incorporate psychophysiological measures as well as self-reported fear. Third, we conducted our research in China, which has a high child pedestrian injury rate and rather chaotic traffic patterns. Generalizability to other cultures and locations is unknown.

5. Conclusion

The role of emotion in children's pedestrian behavior and safety is poorly understood. Our results suggest children report feelings of fear consistent with the risk involved in traffic situations. We also discovered relations between perception of fear and safety of pedestrian decisions. In more challenging traffic situations, children's fear was associated with safer pedestrian decisions. In less challenging situations, fear was associated with street-crossing errors. These findings offer insight into mechanisms that may underlie children's decisions in traffic and may guide development of interventions that incorporate induction of realistic fear of traffic risks into broader lessons on pedestrian safety and efforts to increase children's self-efficacy to handle traffic environments.

Funding

This work was supported by the Key projects of the 13th Five Year Plan for Education Science in Jiangsu Province, China [B-a/2020/01/09] and Major Project of Philosophy and Social Science Research in Jiangsu Province in China [2020SJZDA118]. This research was also partially supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the United States National Institutes of Health under Award Number R01HD088415. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the pupils from Nantong Wushan Primary School and their parents for their participation. We also thank Principal Liu Zhihe and other teachers from the school for their assistance with this study. Finally, we acknowledge Ding Yuye for serving as a research assistant for the study.

References

Ahmadi, A., & Ytterstad, B. (2007). Prevention of self-immolation by community-based intervention. *Burns, 33*(8), 1032–1040.
Albert, R. R., & Dolgin, K. G. (2010). Lasting effects of short-term training on preschoolers' street- crossing behavior. *Accident Analysis and Prevention, 42*, 500–508.
Ampofo-Boateng, K., & Thomson, J. A. (1991). Children's perception of safety and danger on the road. *British Journal of Psychology, 82*(4), 487–505.
Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.
Barton, B. K., & Schwebel, D. C. (2007). The influences of demographics and individual differences on children's selection of risky pedestrian routes. *Journal of Pediatric Psychology, 32*(3), 343–353.
Barton, B. K., Ulrich, T., & Lyday, B. (2011). The roles of gender, age and cognitive development in children's pedestrian route selection. *Child: Care, Health and Development, 38*(2), 280–286.
Bouffard, J. A. (2014). Examining the direct and indirect effects of fear and anger on criminal decision making among known offenders. *International Journal of Offender Therapy and Comparative Criminology, 59*(13), 1385–1408.
Demetre, J. D. (1997). Applying developmental psychology to children's road safety: Problems and prospects. *Journal of Applied Developmental Psychology, 18*(2), 263–270.
Derevensky, J. (1974). What children fear. *McGill Journal of Education, 9*, 77–85.
Dong, X., Peek-Asa, C., Yang, J., Wang, S., Chen, X., Chi, G., & Ramirez, M. (2011). The association of road safety knowledge and risk behaviour with paediatric road traffic injury in Guangzhou, China. *Injury Prevention, 17*(1), 15–20.
Goldenbeld, C., Twisk, D., & Houwing, S. (2008). Effects of persuasive communication and group discussions on acceptability of anti-speeding policies for male and female drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 11*(3), 207–220.
Gullone, E., & Lane, L. (1997). The validity of fear assessment using the Fear Survey Schedule. *Paper presented at the 20th National Conference of the Australian Association of Cognitive and Behaviour Therapy, Brisbane, Australia*.
Institute for Health Metrics and Evaluation, University of Washington. (2019). GBD Compare Viz Hub. Available from<https://vizhub.healthdata.org/gbd-compare/> (Accessed on 25 March 2020).
Job, R. F. S. (1985). Reported attitudes, practices and knowledge in relation to drink-driving: the effects of the introduction of random breath testing in New South Wales. In S. Kaye & G. W. Meier (Eds.), *Alcohol, Drugs and Traffic Safety: Proceedings of the Ninth International Conference on Alcohol, Drugs and Traffic Safety, San Juan, Puerto Rico, 1983*. Washington, DC: National Highway Traffic Safety Administration.
Job, R. F. S. (1988). Effective and ineffective use of fear in health promotion campaigns. *American Journal of Public Health, 78*(2), 163–167.
Kovesdi, C. R., & Barton, B. K. (2013). The role of non-verbal working memory in pedestrian visual search. *Transportation Research Part F, 19*, 31–39.
Koekemoer, K., Gesselleen, M. V., Niekerk, A. V., Govender, R., & As, A. B. V. (2017). Child pedestrian safety knowledge, behaviour and road injury in CapeTown, South Africa. *Accident Analysis and Prevention, 99*(part A), 202–209.
Lewis, I., Watson, B., & White, K. M. (2008). An examination of message-relevant affect in road safety messages: Should road safety advertisements aim to make us feel good or bad? *Transportation Research Part F: Traffic Psychology and Behaviour, 11*(6), 403–417.
Lu, J., Xie, X., & Zhang, R. (2013). Focusing on appraisals: How and why anger and fear influence driving risk perception. *Journal of Safety Research, 45*, 65–73.
Meir, A., Parmet, Y., & Oron-Gilad, T. (2013). Towards understanding child-pedestrians' hazard perception abilities in a mixed reality dynamic environment. *Transportation Research Part F, 20*, 90–107.
Meir, A., Oron-Gilad, T., & Parmet, Y. (2015). Are child-pedestrians able to identify hazardous traffic situations? Measuring their abilities in a virtual reality environment. *Safety Science, 80*, 33–40.
Morrongiello, B. A., & Lasenby-Lessard, J. (2006). Psychological determinants of risk taking by children: An integrative model and implications for interventions. *Injury Prevention, 13*(1), 20–25.
Morrongiello, B. A., Corbett, M., Switzer, J., & Hall, T. (2015). Using a virtual environment to study pedestrian behaviors: How does time pressure affect children's and adults' street crossing behaviors? *Journal of Pediatric Psychology, 40*(7), 697–703.
Morrongiello, B. A., & Matheis, S. (2004). Determinants of children's risk-taking in different–situational contexts: The role of cognitions and emotions in predicting children's decisions. *Applied Developmental Psychology, 25*, 303–326.
Morrongiello, B. A., & Matheis, S. (2007). Understanding children's injury-risk behaviors: The independent contributions of cognitions and emotions. *Journal of Pediatric Psychology, 32*(8), 926–937.
Ollendick, T. H. (1983). Reliability and validity of the Revised Fear Survey Schedule for Children (FSSC-R). *Behaviour Research and Therapy, 21*, 685–692.
Peters, G. J. Y., Ruiter, R. A. C., & Kok, G. (2012). Threatening communication: A critical re-analysis and a revised meta-analytic test of fear appeal theory. *Health Psychology Review, 7*(Suppl. 1), S8–S31.
Poudel-Tandukar, K., Nakahara, S., Ichikawa, M., Poudel, K. C., & Jimba, M. (2007). Risk perception, road behavior, and pedestrian injury among adolescent students in Kathmandu, Nepal. *Injury Prevention, 13*(4), 258–263.
Price, A. E., Corwin, S. J., Friedman, D. B., Laditka, S. B., Colabianchi, N., & Montgomery, K. M. (2011). Older adults' perceptions of physical activity and cognitive health: Implications for health communication. *Health Education & Behavior, 38*(1), 15–24.
Rosenbloom, T., Nemrodov, D., Ben-Eliyahu, A., & Eldror, E. (2008). Fear and danger appraisals of a road-crossing scenario: A developmental perspective. *Accident Analysis and Prevention, 46*(8), 1619–1626.
Ruiter, R. A. C., Abraham, C., & Kok, G. (2001). Scary warnings and rational precautions: A review of the psychology of fear appeals. *Psychology & Health, 16*(6), 613–630.
Schwebel, D. C., Davis, A. L., & O'Neal, E. E. (2012). Child pedestrian injury: A review of behavioral risks and preventive strategies. *American Journal of Lifestyle Medicine, 6*, 292–302.
Schwebel, D. C., & Gaines, J. (2007). Pediatric unintentional injury: Behavioral risk factors and implications for prevention. *Journal of Developmental and Behavioral Pediatrics, 28*, 245–254.
Skinner, B. F. (1969). *Contingencies of reinforcement: A theoretical analysis*. New York: B.F. Skinner Foundation.

Shen, J., McClure, L. A., & Schwebel, D. C. (2015). Relations between temperamental fear and risky pedestrian behavior. *Accident Analysis and Prevention, 80*, 178–184.

Tabibi, Z., & Pfeffer, K. (2007). Finding a safe place to cross the road: The effect of distractors and the role of attention in children's identification of safe and dangerous road-crossing sites. *Infant and Child Development, 16*, 193–206.

Taylor, J. E., Deane, F. P., & Podd, J. V. (2007). Driving fear and driving skills: Comparison between fearful and control samples using standardised on-road assessment. *Behaviour Research and Therapy, 45*, 805–818.

Thomson, J. A. (1997). Developing safe route planning strategies in young child pedestrians. *Journal of Applied Developmental Psychology, 18*, 271–281.

Twisk, D., Vlakveld, W., Mesken, J., Shope, J. T., & Kok, G. (2013). Inexperience and risky decisions of young adolescents, as pedestrians and cyclists, in interactions with lorries, and the effects of competency versus awareness education. *Accident Analysis and Prevention, 55*, 219–225.

Wazana, A., Krueger, P., Raina, P., & Chambers, L. (1997). A review of risk factors for child pedestrian injuries: Are they modifiable? *Injury Prevention, 3*, 295–304.

Whitebread, D., & Neilson, K. (2000). The contribution of visual search strategies to the development of pedestrian skills by 4–11- year-old children. *British Journal of Educational Psychology, 70*(4), 539–557.

Will, K. E., Sabo, C. S., & Porter, B. E. (2009). Evaluation of the Boost' em in the Back Seat Program: Using fear and efficacy to increase booster seat use. *Accident Analysis and Prevention, 41*, 57–65.

World Health Organization. (2018). Global status report on road safety 2018. Available from <www.who.int/violence_injury_prevention/road_safety_status/2015/en/> (accessed on 8 April 2020).

Huarong Wang Ph.D., is an Associate Professor in the Department of Traffic Psychology at Nantong University. Her research interests are road traffic safety and impact mechanism for children and adolescents.

Casie Morgan, BA, is a doctoral candidate in the University of Alabama at Birmingham Medical/Clinical Psychology program. She is interested especially in links between behavior problems, temperament, and injury.

Dongqian Li, BA, is a master candidate in Nantong University. She is interested in the traffic psychology and child safety.

Rong Huang, BA, is a undergraduate in Nantong University. She is interested in the developmental psychology and child safety

David C Schwebel Ph.D., is a Professor in the Department of Psychology at University of Alabama at Birmingham. His primary research areas of interest is understanding and preventing unintentional child injuries.

Journal of Safety Research 76 (2021) 73–82

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Data mining approach to model bus crash severity in Australia

Seyed Alireza Samerei [a], Kayvan Aghabayk [a,*], Amin Mohammadi [b], Nirajan Shiwakoti [c]

[a] School of Civil Engineering, College of Engineering, University of Tehran, Tehran, Iran
[b] Mianeh Technical and Engineering Faculty, University of Tabriz, Tabriz, Iran
[c] School of Engineering, RMIT University, Melbourne, Australia

A R T I C L E   I N F O

Article history:
Received 6 June 2020
Received in revised form 20 August 2020
Accepted 4 December 2020
Available online 20 December 2020

Keywords:
Bus involved crashes
Crash severity
Public transport
Clustering
Association rule discovery

A B S T R A C T

Introduction: Buses are different vehicles in terms of dimensions, maneuverability, and driver's vision. Although bus traveling is a safe mode to travel, the number of annual bus crashes cannot be neglected. Moreover, limited studies have been conducted on the bus involved in fatal crashes. Therefore, identification of the contributing factors in the bus involved fatal crashes can reduce the risk of fatality. *Method:* Data set of bus involved crashes in the State of Victoria, Australia was analyzed over the period of 2006–2019. Clustering of crash data was accomplished by dividing them into homogeneous categories, and by implementing association rules discovery on the clusters, the factors affecting fatality in bus involved crashes were extracted. *Results:* Clustering results show bus crashes with all vehicles except motor vehicles and weekend crashes have a high rate of fatality. According to the association rule discovery findings, the factors that increase the risk of bus crashes with non-motor vehicles are: old bus driver, collision with pedestrians at signalized intersections, and the presence of vulnerable road users. Likewise, factors that increase the risk of fatality in bus involved crashes on weekends are: darkness of roads in high-speed zones, pedestrian presence at highways, bus crashes with passenger car by a female bus driver, and the occurrence of multi-vehicle crashes in high-speed zones. *Practical Applications:* The study provides a sequential pattern of factors, named rules that lead to fatality in bus involved crashes. By eliminating or improving one or all of the factors involved in rules, fatal bus crashes may be prevented. The recommendations to reduce fatality in bus crashes are: observing safe distances with the buses, using road safety campaigns to reduce pedestrians' distracted behavior, improving the lighting conditions, implementing speed bumps and rumble strips in high-speed zones, installing pedestrian detection systems on buses and setting special bus lanes in crowded areas.

1. Introduction

Road traffic injuries are one of the top 10 causes of fatality in the world (WHO, 2019). It is estimated that about 1.35 million people are killed in traffic crashes worldwide annually (WHO, 2019). Meanwhile, public bus transportation is considered as one of the most common and safe modes of travel in developed countries and its use among people is increasing (Barua & Tay, 2010). Statistically, the number of bus crashes is less than 1% of the total traffic crashes, however the high passenger carrying capacity of this vehicle compared to passenger cars increases the amount of financial and human losses caused to it (Chimba et al., 2010). Although bus travel is one of the safest modes of travel, the number of bus crashes cannot be neglected (Chimba et al., 2010). In some parts of the world, the number of bus crashes has increased and the knowledge about bus safety is usually less than the safety of passenger cars (Feng et al., 2016). Although the number of people killed in bus crashes in Australia in the first half of the last decade (2009–2018) has been declining, there has been a slight increase in the number of people killed in bus crashes over the past five years (BITRE, 2019). Existing studies have focused mainly on the analysis of risk factors associated with the likelihood of bus crashes and the bus driver errors (Kaplan & Prato, 2012). Despite the great interest in bus safety, bus crashes have received less attention and many fundamental questions of bus crashes remain unanswered (af Wåhlberg, 2004; Chimba et al., 2010). It is important to identify the factors and the pattern of bus crashes that leads to fatality.

Combined models can detect basic hidden patterns in crash data and discover the effect of factors on the severity of the injury (Sun et al., 2019). Due to the inherent heterogeneity of crash data, which is usually due to the lack of reporting of some of the factors

* Corresponding author.
E-mail addresses: alireza.samerei@ut.ac.ir (S.A. Samerei), kayvan.aghabayk@ut.ac.ir (K. Aghabayk), amin.mohammadi@tabrizu.ac.ir (A. Mohammadi), nirajan.shiwakoti@rmit.edu.au (N. Shiwakoti).

https://doi.org/10.1016/j.jsr.2020.12.004
0022-4375/

involved in the crash and the non-uniform distribution of factors affecting the crashes, clustering of crash records is necessary (Li & Fan, 2019; Sun et al., 2019). In addition, segmentation of the database leads to a better understanding of the complex relationships between the severity of the crash and the contribution of geometric, environmental, and demographic factors of drivers. Recently, some researchers have used clustering techniques to create data clusters that were later used as inputs in subsequent analyzes (Alikhani et al., 2013; Kashani & Besharati, 2017; Mohamed et al., 2013). In terms of analytical methods, various regression models have been used to model the occurrence of crashes (Lord & Mannering, 2010) and the severity of road users' injuries (Kaplan & Prato, 2012). Data mining techniques have been introduced and used in recent years to analyze large data sets of traffic crashes. Data mining involves several parametric and non-parametric techniques that can be used to analyze large amounts of data and extract hidden patterns (Kumar & Toshniwal, 2016). Over the past decade, non-parametric data mining methods, including association rules discovery, have been used to analyze crash data (Besharati & Tavakoli Kashani, 2018; Montella, 2011; Pande & Abdel-Aty, 2009; Weng et al., 2016).

In summary, limited studies have been conducted on the bus involved fatal crashes, especially identifying the factors and the pattern of bus crashes that leads to fatality. Therefore, this study aims to identify the contributing factors in the bus involved fatal crashes that can reduce the risk of fatality. Since a crash is caused by the chain of factors, the severity of the crash injury can be reduced by eliminating one or all of them. Unlike other models that measure the impact of crash factors on the severity of a crash individually, this study uses association rules discovery to discover chains of effective factors that lead to a fatality in bus involved crashes. In the present research, factors related to the environment, humans, and vehicles have been considered. For this purpose, bus crashes data in the state of Victoria, Australia during 2006–2019 have been used. The data included 2.7% of fatal crashes, 34% of severe injuries, and the rest classified as the other types of injuries (VicRoads, 2019). In addition, the rules provided by this model introduce the pattern of occurrence of a specific severity of the bus involved crash, which is statistically reliable. Knowing these patterns and the parameters involved in them will assist road authorities to develop appropriate strategies to reduce the number of casualties and injuries caused by the collision with the bus.

The paper is organized as follows. Section 2 reviews the previous studies on the severity of injuries in bus crashes. Section 3 describes the methodology adopted for this study that includes an explanation of the two-step clustering and association rules discovery models. Section 4 is devoted to bus crash data used in this study. It is followed by Section 5 that presents the discussion of results from the models. Finally, in Section 6, the conclusions of this study are presented.

2. Literature review

The main factors involved in a crash are often described under three components: driver, vehicle, and environment (Goh et al., 2014a). It should be noted that the safety risk associated with various human, physical, and environmental aspects of bus performance is complex (Strathman et al., 2010). Some of the main problems in this area are: differences in reports published by hospitals and police, inaccuracies in the report of property damage only (PDO) crashes (which is the most common level of bus crashes), insufficient data about the type and severity of injuries and the lack of an acceptable classification of bus crashes (Albertsson & Falkmer, 2005). In addition, buses may be directly or indirectly involved in crashes (Brenac & Clabaux, 2005). More than 40% of the bus crashes were rear-ended, while 80% of the crashes occurred when the bus stopped at the station and was hit by another vehicle (Jovanis et al., 1991). Rollover is a common bus crash, especially in suburban areas (Xiaoyun et al., 2019) where the risk of fatality is five times higher than other types of bus crashes (Martínez et al., 2003). Studies show that most bus crashes are caused by a violation of the right of way, control loss of the vehicle, reckless and high-speed driving. In addition, the close proximity of the bus to other vehicles is one of the main reasons for bus crashes (Chu, 2014; Li et al., 2015; Tseng, 2012) and bus crashes with the fixed object are common (Li et al., 2012).

The number of urban bus crashes is higher, while the suburban bus crashes are more severe (Albertsson & Falkmer, 2005). Bus involved crashes that occur at intersections also cause more serious injuries (Kaplan & Prato, 2012). Increasing the number of lanes and the volume of traffic on each lane increases the likelihood of bus crash, while increasing the width of lanes and midways reduces the likelihood of bus crash (Chimba et al., 2010). Further, the movement of the bus on the inner lanes increases the likelihood of colliding with the buses (Chimba et al., 2010). Likewise, increasing the width of the lane and shoulder and reducing the volume of traffic in each lane will reduce the severity of injuries (Chimba et al., 2010). The implementation of bus priority in addressing maneuverability in Melbourne, Australia has shown a significant reduction in the ratio of the bus involved crashes (Goh et al., 2014c). Although the number of bus fatalities and injuries in Hong Kong has decreased with the implementation of bus rapid transit systems, other users have been exposed to more severe injuries while using exclusive bus lane (Tse et al., 2014). Clabaux et al. investigated the risk of a bus crash with other users at bus rapid transit system and found that due to the significant speed difference between buses and motorcyclists (powered two-wheeler drivers) using exclusive bus lanes, the risk of more severe injuries to motorcyclists was significantly increased (Clabaux et al., 2014).

Bus involved crashes can cause serious damage due to features such as heavyweight, large size, and limitations such as limited maneuverability (Yoon et al., 2017). High-deck buses, which have a higher center of gravity and travel at high speeds, increase the risk of serious crashes and casualties when losing control of vehicles (Chu, 2014).

Inexperienced and young drivers, as well as older drivers, increase the probability of a crash (Kaplan & Prato, 2012). Zegeer et al. found no association between bus crashes and gender and the age of bus drivers (Zegeer et al., 1993). Kaplan and Prato concluded that young drivers were more risky to be involved in crashes and that male drivers increased the probability of crashes with less severity (Kaplan & Prato, 2012). Goh et al. showed that bus drivers over the age of 60, those with less than two years of driving experience and drivers with experience of crashes increase the likelihood of bus crashes (Goh et al., 2014a). Moreover, the presence of vulnerable road users, such as pedestrians, directly affects the severity of bus crashes (Prato & Kaplan, 2014).

Crash severity increases when a crash occurs over the weekend, at two-way streets, and the presence of pedestrians and other vulnerable users (Barua & Tay, 2010). Research in Ghana (Sam et al., 2018) also showed that the weekend, lack of median, night time, poor road conditions (curved, wet and rough roads) and drunk driving have increased the severity of the crashes, and on the other hand, the lack of road shoulders, crashes at intersections and traffic control systems reduce their severity.

3. Methodology

Our objective is to investigate the relationship between the bus involved crash severity and chains of the effective factors (environment, humans, and vehicles) that lead to a fatality in bus involved crashes. Due to the large number of parameters involved in the crashes and inherent heterogeneity of the crash data and the non-uniform distribution of factors affecting the crash, clustering is necessary. For this reason, in the first step, the crash data was divided into more homogeneous groups so that the severity of crashes in each cluster that had unique conditions could be examined more carefully. In the second step, using the association rules discovery method, a chain of contributing factors leading to the fatality of the passengers in the collision with the bus was discovered. The rules presented by this model also introduced the pattern of occurrence leading to the fatality in bus crashes, which is statistically valid. The details of the models are described below.

3.1. Two-step clustering

In this study, the two-step clustering algorithm is used to divide crash data into similar categories (Chiu et al., 2001). This algorithm uses a hierarchical clustering method called BIRCH (Zhang et al., 1996). The mentioned algorithm consists of two steps. In the first step, all records are scanned and the location of the record density is determined. In the second step, the specified density location of records is received as the input and a hierarchical clustering algorithm based on log-likelihood distance is used to group them into the desired number of clusters.

The basic strategy for determining the most appropriate number of clusters is to calculate Bayesian Information Criterion (BIC). In the first step, the BIC is calculated for each number of clusters in a specific domain to obtain an initial estimate of the number of clusters. BIC for a model with k cluster is:

$$\mathrm{BIC}(k) = -2\mathrm{LL}(k) + r(k)\log N \tag{1}$$

$\mathrm{LL}(k)$ is a log-likelihood function for the model with k cluster, $r(k)$ is the number of independent parameters, and N is the total number of records in the data set. The BIC change ratio in each merge relative to the first merge then determines the initial estimate. $\mathrm{dBIC}(k)$ is the difference between BIC between the model with k cluster and $(k+1)$ cluster. Then the ratio of BIC changes according to the changes of the two clusters as follows:

$$\mathrm{RBK}(k) = \mathrm{dBIC}(k)/\mathrm{dBIC}(1) \tag{2}$$

The ratio of measuring the distance $R(k)$ for kth cluster is as follows:

$$\mathrm{R}(k) = \mathrm{RBK}(k)/\mathrm{RBK}(k+1) \tag{3}$$

Finally, the two large ratios of R is compared. If the largest is more than 1.15 times the second largest R, then the model with the highest R ratio is selected as the optimal number of clusters (Hamid & Abawajy, 2014). It should be noted that the final solution might depend on the order of the records in the database. Therefore, these files are ordered randomly to ensure that the data sequence does not have a significant effect on the final solution. The two-step cluster analysis was performed in the present study using SPSS software v.25.

3.2. Association rules discovery

Association rules discovery, also known as market basket analysis, is a prevalent data mining technique for discovering relationships between different attributes (Besharati & Tavakoli Kashani, 2018). This method is well suited for unbalance data such as crash data. In other words, this model is based on the relative frequency of occurrence of the sets of items alone and in combination with each other. This method is used to find the crash factors that occur more frequently than when they are statistically independent.

Association rules discovery was first proposed by Agrawal and Srikant (1994). In this study, the apriori algorithm proposed by Agrawal and Srikant (1994) was used to discover association rules. This algorithm uses multiple criteria with predefined threshold values to identify frequent item sets and create association rules. In this study, the rules are determined by the values of "Support," "Confidence," and "Lift." These values are defined as follows:

Support of a rule is the percentage of the entire data-set covered by the rule and is expressed as:

$$\mathrm{support}(\mathrm{A} \rightarrow \mathrm{B}) = \#(A \cap B)/N \tag{4}$$

where N is the total number of crashes, (A∩B) is the number of crashes in which both A (antecedent) and B (consequent) factors have co-occurred.

Confidence of a rule is the conditional probability of occurrence of the consequent, given that the antecedents have occurred and is expressed as:

$$\mathrm{Confidence} = \mathrm{support}(\mathrm{A} \rightarrow \mathrm{B})/\mathrm{support}(\mathrm{A}) \tag{5}$$

Lift of a rule is the ratio of the confidence of the rule and its expected value. The lift of an association rule (A → B) can be calculated as:

$$\mathrm{Lift} = \mathrm{support}(\mathrm{A} \rightarrow \mathrm{B})/(\mathrm{support}(\mathrm{A}) \times \mathrm{support}(\mathrm{B})) \tag{6}$$

Lift indicates frequency of simultaneous occurrence of antecedent and consequent (i.e. the effect of occurrence of antecedents on the conditional probability of occurrence of the consequent). The value of lift as equal to one, represents that having the antecedent(s) does not make a lot of difference in the probability of having the consequent; while a lift value greater than one shows a positive interdependence between antecedent and consequent. The higher the Lift value, the probability of occurrence of the antecedent and consequent in an event is not by chance, and the greater the interdependence between them.

Any rule with $n+1$ items is valid if by adding a variable, the Lift increases sufficiently. By applying LIC criterion (as defined in Equation (7)), Lift increase is checked. Rules with one antecedent are used as a base of validity and the start point. Other rules with more antecedent will be valid, if the minimum LIC threshold is reached (1.05) (López et al., 2014; Montella et al., 2011, 2012, 2020). LIC is determined as below:

$$\mathrm{LIC} = \mathrm{Lift}(A_{n+1})/\mathrm{Lift}(A_n) \tag{7}$$

where A_n is the antecedent of rule with n items, and A_{n+1} is the antecedent of rule with n + 1 items.

Due to the fact that the subject of this study is fatal severity and that it constitutes less percentage than other severity of crashes, so the minimum threshold values for support (S), confidence (C), lift (L) and LIC are considered, 1%, 5%, 1.2 and 1.05 respectively. These values were obtained after several trial and error experiments (Das et al., 2019) and considering the threshold values taken in other studies (Besharati & Tavakoli Kashani, 2018; Montella et al., 2012, 2020). It should be noted that although the amount of support in a crash database may be low for some association rules due to its rarity in crashes, it has a strong significant relationship. This means that the lift criterion is more important in determining how strong an association rule is than the other two criteria (Pande & Abdel-Aty, 2009).

4. Data

In this study, records of bus crashes in Victoria, Australia between 2006 and 2019 have been used (VicRoads, 2019). All roads in the Victoria State in Australia including metropolitan Melbourne roads and rural roads were examined, which are divided into six main regions: Metropolitan South East Region, Eastern Region, South Western Region, Northern Region, and North Eastern Region (VicRoads, 2019). In order to examine more accurately the factors affecting the severity of the bus crashes, all the records in which the bus was involved were examined. After deleting the missing information records, 1,705 bus crashes were finally analyzed. Possible factors influencing the occurrence of the crashes, including driver characteristics, environmental characteristics, geometry and traffic characteristics, and crash characteristics, have been considered. Moreover, heavy vehicle (heavy commercial vehicle, light commercial vehicle, rigid truck) and vulnerable users (pedestrians, bicycle, motor cycle, motor scooter, quad bike, and moped) were also investigated in bus involved crashes. It is to be noted that some other factors may influence the occurrence of crashes such as annual average daily traffic (AADT) and the width of the route on the severity of the injury. However, these factors could not be considered in the present study as they were not available in the crash database.

Australian passenger transport buses, based on their size and weight, are classified to three groups: Light Rigid (4.5–8 tons), Medium Rigid (more than 8 tons and has no more than 2 axles) and Heavy Rigid (a bus consisting of more than one rigid section which are connected to one another). According to crash reports in Victoria State in 2006–2019, the buses under study include 48 different brands. The most frequent brands in crashes include: Toyota (10%), Mercedes B (15%), Volvo (22%), Scandia (23%), others (30%) with capacity seating of 15 to 65 passengers and the weight of 3 to 25 tons which are made in 1980–2017 (VicRoads, 2019). In this study Van, Minibus, and Wagon have not been considered.

Based on the distribution of bus crash severity for each variable which is shown in Table 1, bus overturned is a rare crash type but has a high rate of fatality (7.2%) and has been observed in other studies (Seyedi et al., 2019). Also, the fatality rate of bus crashes in highways and high-speed zones are 6.7% and 3.3% respectively; and in comparison to the average fatality rate, has a considerable increase (2.7%). Bus crashes at the weekend, at nights, and by inexperienced drivers have high rate of fatality in Australia (5.5%, 5.5%, and 5.2%, respectively).

According to Australian law, the legal age to get a driving license is 18. If drivers want to get bus license of Light and Medium Rigid buses, should have at least one year driving experience and for Heavy Rigid buses, two years driving experience is necessary. The driver's age in Australia is between 19 to 78 years old (mean = 47.5, SD = 14). Although the crash rate of drivers under 20 is very low, the fatality rate in this group is considerable (8.8%) and it is noted in similar studies (Goh et al., 2014a). Distribution of bus crash severity versus other variables is represented in Table 1.

The term severity, which is used in different parts of this paper, means the severity of injuries to road users due to crashes and includes 4 levels of injuries, which are: fatality, serious injury, others injury, and no injuries. People who die at the scene of a crash or up to 30 days after the crash are in the fatality group. People who have serious injuries and are sent to the hospital are in the group of serious injuries and people who have other injuries such as bruising, contusions, and pain are in the group of others injury. According to the Table 1, 2.7% of bus crashes constitute the fatality group. Serious injuries occurred in 34% of bus crashes while others injury occurred in 63% of bus crashes.

It is to be noted that while in some studies the variables fatality and serious injury are considered as one variable (Montella et al., 2020; Nasri & Aghabayk, 2020), in many other studies, these variables have been considered separately (Besharati & Tavakoli Kashani, 2018; Kaplan & Prato, 2012; Kashani et al., 2014; Montella et al., 2012; Sam et al., 2018). By following the latter approach, the fatality and serious injury considered separately and the fatality was defined as target variable.

5. Results and discussion

5.1. Clustering analysis

The BIC change for each potential solution (k ranging from 2 to 3) is given in Table 2. The two largest values of R (k) are: R (2) = 1.358 and R (3) = 2.315. Therefore, the R ratio is equal to R (3)/R (2) = 2.315 / 1.358 = 1.70, which is greater than the threshold of 1.15. Thus, the number of cluster 3 was selected as the optimal number of clusters.

In the next step, the clusters were analyzed and named based on their variable distribution. In other words, the variables that allow the distinction between clusters were used to describe each cluster. All variables were examined and finally, two variables were selected to describe the clusters. Based on the single variable distribution for the variables in Table 3, the following three clusters are listed as follows:

- Cluster 1: Collision with motor vehicles at weekdays
- Cluster 2: All types of collisions except collisions with motor vehicles at weekdays
- Cluster 3: Weekend

By explaining how the variables are distributed between the 3 clusters, the crash pattern of each cluster can be defined. As shown in Table 3, 100% of Cluster 1 crashes occurred on weekdays. In addition, all the records of Cluster 1 crashes in terms of the type of collision are related to bus crashes with motor vehicles. Similar to Cluster 1, all Cluster 2 crashes occurred on weekdays and include all collisions except collisions with motor vehicles. In cluster 3, 100% of bus crashes occurred on weekends.

The general description of the clusters is given in Table 4. It can be observed that more than half of bus crashes (55%) are related to collisions with motor vehicles on weekdays (cluster 1). In this cluster, the fatality rate is 1.6%, which is lower than the average fatality rate (2.7%). According to literature (Damsere-Derry et al., 2017), bus and car occupants have a protected shell and therefore, they have more relative safety features as compared to vulnerable road users, including pedestrians. The presence of safety features in passenger cars, including seat belts and the relative protection of the occupants inside the car and the bus compared to vulnerable road users, can be the reason for the low rate of fatalities in this cluster (Damsere-Derry et al., 2017). In addition, all the records of crashes in this cluster occurred in the working days of the week and show the different pattern of crashes in the days of the week. This observation is consistent with the findings from other studies too (Feng et al., 2016).

Cluster 2 deals with all bus collisions except for collisions with motor vehicles on weekdays. The cluster accounts for 27% of all crashes and has a 3.2% fatality rate, which is higher than the average (2.7%). According to the distribution of variables of Cluster 2 in Table 3, the highest number of crashes in this cluster is related to the collision with pedestrians or fall from or in moving vehicle. The high rate of fatalities in bus crashes in the presence of vulnerable users such as pedestrians seems reasonable, and according to other

Table 1
Variable description for all bus involved crashes.

Variable	Description	Fatality	Serious injury	Other injuries	Total
Bus involved crash severity		46(2.7%)	585(34.3%)	1074(63%)	1075(100%)
Definition of classified accidents (DCA)	Collision with a fixed object	3(5.1%)	18(30.5%)	38(64.4%)	59(3.5%)
	Collision with some other object	0(0%)	2(13.3%)	13(86.7%)	15(0.9%)
	Collision with motor vehicle	21(1.9%)	372(33.5%)	718(64.6%)	1111(65.2%)
	Fall from or in moving vehicle	1(0.5%)	65(30.8%)	145(68.7%)	211(12.4%)
	No collision and no object struck	0(0%)	11(21.6%)	40(78.4%)	51(3%)
	Struck animal	1(25%)	1(25%)	2(50%)	4(0.2%)
	Struck pedestrian	0(0%)	1(20%)	4(80%)	5(0.3%)
	Bus overturned (no collision)	17(7.2%)	108(45.8%)	111(47%)	236(13.8%)
Traffic control	Give way sign	1(1.3%)	29(37.7%)	47(61%)	77(4.5%)
	No control	29(2.6%)	360(32.5%)	717(64.8%)	1106(64.9%)
	Ped. Crossing	1(3.6%)	13(46.4%)	14(50%)	28(1.6%)
	Ped. Lights	0(0%)	2(40%)	3(60%)	5(0.3%)
	Roundabout	1(2.2%)	14(31.1%)	30(66.7%)	45(2.6%)
	School flags	0(0%)	1(16.7%)	5(83.3%)	6(0.4%)
	Stop sign	1(4%)	11(44%)	13(52%)	25(1.5%)
	Stop-go lights	12(3%)	154(38.2%)	237(58.8%)	403(23.6%)
	Others	1(10%)	1(10%)	8(80%)	10(0.6%)
Road geometry	Cross intersection	12(2.4%)	185(37.5%)	296(60%)	493(28.9%)
	Multiple intersection	3(7.1%)	16(38.1%)	23(54.8%)	42(2.5%)
	Not at intersection	23(3%)	261(33.8%)	488(63.2%)	772(45.3%)
	T intersection	8(2%)	123(31.1%)	264(66.8%)	395(23.2%)
	Y intersection	0(0%)	0(0%)	3(100%)	3(0.2%)
Road type	Avenue	0(0%)	7(20%)	28(80%)	35(2.1%)
	Freeway	2(3.9%)	17(33.3%)	32(62.7%)	51(3%)
	Highway	11(6.7%)	60(36.8%)	92(56.4%)	163(9.6%)
	Road	23(2.9%)	276(35.1%)	488(62%)	787(46.2%)
	Street	7(1.3%)	164(31.4%)	352(67.3%)	523(30.7%)
	Others	3(2.1%)	61(41.8%)	82(56.2%)	146(8.6%)
Speed limit	Over50	41(3.3%)	436(34.8%)	776(61.9%)	1253(73.5%)
	UP to50	5(1.1%)	149(33%)	298(65.9%)	452(26.5%)
Light condition	Darkness	12(5.5%)	98(44.7%)	109(49.8%)	219(12.8%)
	Day	29(2.1%)	452(33.1%)	884(64.8%)	1365(80.1%)
	Dusk/Dawn	5(4.1%)	35(28.9%)	81(66.9%)	121(7.1%)
Road Surface Type	Gravel	0(0%)	14(46.7%)	16(53.3%)	30(1.8%)
	Paved	46(2.8%)	568(34.2%)	1047(63%)	1661(97.4%)
	Unpaved	0(0%)	3(25%)	9(75%)	12(0.7%)
Weather condition	Clear	40(2.6%)	513(33.6%)	976(63.8%)	1529(89.7%)
	Fog/Snowing	0(0%)	4(30.8%)	9(69.2%)	13(0.8%)
	Raining	3(2.1%)	64(44.8%)	76(53.1%)	143(8.4%)
	Strong winds/Dust	3(15%)	4(20%)	13(65%)	20(1.2%)
Surface condition	Dry	38(2.6%)	494(33.5%)	943(63.9%)	1475(86.5%)
	Icy/Snowy	1(16.7%)	1(16.7%)	4(66.7%)	6(0.4%)
	Wet/Muddy	7(3.1%)	90(40.2%)	127(56.7%)	224(13.1%)
Heavy vehicle presence	Heavy vehicle exist	2(2.6%)	30(39.5%)	44(57.9%)	76(4.5%)
	Heavy vehicle not exist	44(2.7%)	555(34.1%)	1030(63.2%)	1629(95.5%)
Vulnerable presence	Vulnerable exist	26(5.5%)	188(40%)	256(54.5%)	470(27.6%)
	Vulnerable not exist	20(1.6%)	397(32.1%)	818(66.2%)	1235(72.4%)
No. vehicle involved	1 vehicle	23(4.6%)	177(35.5%)	298(59.8%)	498(29.2%)
	More than 1 vehicle	23(1.9%)	408(33.8%)	776(64.3%)	1207(70.8%)
Day of the week	Weekday	30(2.1%)	463(33.1%)	905(64.7%)	1398(82%)
	Weekend	16(5.2%)	122(39.7%)	169(55%)	307(18%)
Driver gender	Female	9(4.5%)	65(32.3%)	127(63.2%)	201(11.8%)
	Male	37(2.5%)	520(34.6%)	947(63%)	1504(88.2%)
Driver age group	Teen driver (under 20)	5(8.8%)	22(38.6%)	30(52.6%)	57(3.3%)
	Young driver (20–40)	9(1.8%)	156(31.7%)	327(66.5%)	492(28.9%)
	Mid age driver (41–60)	21(2.5%)	289(34.6%)	526(62.9%)	836(49%)
	Old driver (over 60)	11(3.4%)	118(36.9%)	191(59.7%)	320(18.8%)

Table 2
Changes in BIC for k ranging from 2 to 3.

Number of Clusters	Schwarz's Bayesian Criterion (BIC)	BIC Change[a]	Ratio of BIC Changes[b]	Ratio of Distance Measures[c]
1	5571.985			
2	3682.314	−1889.671	1.000	1.358
3	2308.054	−1374.260	0.727	2.315

[a] The changes are from the previous number of clusters in the table.
[b] The ratios of changes are relative to the change for the two cluster solution.
[c] The ratios of distance measures are based on the current number of clusters against the previous number of clusters.

Table 3
Summary of univariate distributions for the variables in each cluster.

Variable	Level	Freq./Percent	Cluster 1	Cluster 2	Cluster 3
Day of Week	Weekday	Frequency	937	461	0
		Percent	**100.0%**	**100.0%**	0.0%
	Weekend	Frequency	0	0	307
		Percent	0.0%	0.0%	**100.0%**
Crash Type Description	Collision with a fixed object	Frequency	0	44	15
		Percent	0.0%	9.5%	4.9%
	collision with some other object	Frequency	0	13	2
		Percent	0.0%	2.8%	0.7%
	Collision with motor vehicle	Frequency	937	0	174
		Percent	**100.0%**	**0.0%**	56.7%
	Fall from or in moving vehicle	Frequency	0	158	53
		Percent	0.0%	34.3%	17.3%
	No collision and no object struck	Frequency	0	42	9
		Percent	0.0%	9.1%	2.9%
	Other accident	Frequency	0	4	0
		Percent	0.0%	0.9%	0.0%
	Struck animal	Frequency	0	3	2
		Percent	0.0%	0.7%	0.7%
	Struck Pedestrian	Frequency	0	187	49
		Percent	0.0%	40.6%	16.0%
	Vehicle overturned (no collision)	Frequency	0	10	3
		Percent	0.0%	2.2%	1.0%

Table 4
Cluster descriptions.

Cluster No.	Cluster description	Share of each cluster from entire database	Percentage of fatal crashes in each cluster	Serious injury	Other injury	Not injured
1	Collision with motor vehicles at weekdays	55.0% (937)	1.6%)15)	32.4% (304)	66.0% (618)	0.0%(0)
2	All types of collisions except collisions with motor vehicles at weekdays	27.0% (461)	3.2% (15)	34.5% (159)	62.3% (287)	0.0%(0)
3	Weekend	18.0% (307)	5.2% (16)	39.7% (122)	55.1% (169)	0.0%(0)
Total		100.0% (1705)	2.7%(46)	34.3% (585)	63.0% (1074)	0.0%(0)

studies, the presence of pedestrians increases the rate of fatalities in bus involved crashes (Sam et al., 2018).

Cluster 3 crashes are related to bus collisions on the weekends, accounting for 18% of all bus involved crashes. The fatality rate of this cluster is 5.2%, which is almost twice the average fatality rate of bus involved crashes (2.7%). Changes in the traffic and behavioral patterns of drivers and other users on weekends can be the cause of high bus crash fatalities on weekends. Other studies have pointed to a different pattern of vehicle crashes over the weekend than on weekdays (Yu & Abdel-Aty, 2013). Crashes on weekends have also been reported in other studies, with increase in the rate of fatalities in bus involved crashes (Barua & Tay, 2010; Sam et al., 2018). The high severity of bus crash fatalities over the weekend could be due to driving under the influence of alcohol or drugs by road users on weekends (Schepens et al., 1998).

5.2. Association rules

It is worth noting that association rules are valid only if all antecedents coincide. In other words, if any of the antecedents can be removed; there will be no more association rules. Association rules for bus collision with motor vehicles on weekdays (cluster 1) are shown in Table 5, which satisfies the LIC criterion. The first rule of Table 5 shows that bus crashes on suburban roads will increase the probability of fatalities by 1.34 times. In other studies, the number of casualties on suburban roads has been significant too (Albertsson & Falkmer, 2005), and the severity of bus crashes on roadways has been more severe than on roundabouts and bus stations (Nasri & Aghabayk, 2020). According to Rule 2, if the speed limit on suburban roads is above 50 km/h, the probability of fatalities is 1.40 times more than the average. These results are consistent with the results of the direct effect of speed limit on the severity of bus crash injuries (Chimba et al., 2010). Additionally, according to Rule 3, if the speed limit on suburban roads is above 50 km/h and the bus collides with more than one vehicle, the probability of fatalities is 1.48 times more than the average. In addition, the impact of the number of vehicles involved on the severity of bus crashes has been investigated and emphasized (Feng et al., 2016).

Association rules apply to all bus collisions except for collisions with motor vehicles on weekdays (cluster 2) as shown in Table 6, which satisfies the LIC criterion. The fatality rate in this cluster is higher than the average bus crash fatality rate (3.2%). According to Rule 1, older bus drivers increase the probability of fatalities

Table 5
List of interesting rules identified for data in cluster 1 (Collision with motor vehicles at weekdays).

LIC	Lift	C%	S%	Consequent	Antecedent	ID
n.a.	1.34	5.1	1.8	Fatality	Road type = “Road”	1
1.05	1.40	6.2	1.2	Fatality	Road type = “Road” & Speed limit = “Over50”	2
1.06	1.48	6.2	1.2	Fatality	Road type = “Road” & Speed limit = “Over50” & No. vehicle = “more than 1veh”	3

Table 6
List of interesting rules identified for data in cluster 2 (All types of collisions except collisions with motor vehicles at weekdays).

LIC	Lift	C%	S%	Consequent	Antecedent	ID
n.a.	1.87	6.0	1.0	Fatality	Driver age group = "Old Driver"	1
1.06	1.98	6.3	1.0	Fatality	Driver age group = "Old Driver" & Driver gender = "Male"	2
n.a.	1.57	5.1	2.6	Fatality	Vulnerable presence = "Vulnerable exist"	3
1.15	1.81	5.9	2.3	Fatality	Vulnerable presence = "Vulnerable exist" & DCA = "Struck Pedestrian"	4
1.09	1.98	6.3	1.1	Fatality	Vulnerable presence = "Vulnerable exist" & DCA = "Struck Pedestrian" & Traffic control = "No control"	5
2.12	3.34	10.8	1.0	Fatality	Vulnerable presence = "Vulnerable exist" & Driver age group = "Old Driver"	6
n.a.	1.89	6.1	1.0	Fatality	Traffic control = "Stop-go lights"	7
1.32	2.51	8.1	1.0	Fatality	Traffic control = "Stop-go lights" & DCA = "Struck Pedestrian"	8

by 1.87 times. This may be due to increase in reaction time among older drivers. It has been reported that driving performance changes steadily across age groups with reaction time and variability of driving performance increased progressively between the ages of 20 and 80 (Svetina, 2016). Other studies have also emphasized the effect of age on the severity of bus crashes (Feng et al., 2016) and that the older age of the bus driver increases the likelihood of fatalities (Prato & Kaplan, 2014). According to Rule 2, old bus drivers with male gender increase the probability of fatalities by 1.98 times. In previous studies as well, the gender of the bus driver is affected by the severity of the crashes (Feng et al., 2016).

According to Rule 3, the presence of vulnerable users in bus involved crashes will increase the risk of the fatalities by 1.57 times than the average. According to previous studies, bus collisions with vulnerable road users increase the risk of crash fatalities. The collision of buses and minibuses with pedestrians (Sam et al., 2018) and cyclists or motorcyclists using special bus lanes (Clabaux et al., 2014) increases the severity of bus involved crashes. Also, according to Rule 4 and 5, if the vulnerable user is a pedestrian and collision occurs in the absence of special traffic control, the probability of fatality increases to 1.81 and 1.98 times, which is consistent with the results of direct effects of the presence of pedestrians on the severity of bus crashes (Prato & Kaplan, 2014).

The highlight of this study is the investigation of the concurrent occurrence of older age of drivers and the presence of vulnerable road users in bus crashes. According to Rule 6, the simultaneous occurrence of old bus driver and vulnerable users, increases the probability of fatalities by 3.34 times, which is much more risky than other rules. According to studies (Kaplan & Prato, 2012), driving over the age of 65 increases the risk of bus crashes. Besides, according to studies conducted on city buses, the presence of a vulnerable road users in bus involved crashes is directly related to the increase in the severity of injuries (Nasri & Aghabayk, 2020).

According to Rule 7, bus crashes at signalized intersection, increases the probability of fatalities by 1.89 times. Intersections are usually the most dangerous traffic sites for pedestrians (Wei et al., 2014) and other vulnerable users like motorcyclists (Vajari et al., 2020), which can be due to the complex bus maneuvers at intersections. Also, according to Rule 8, a direct collision of a bus and pedestrians at a signalized intersection, increases the probability of fatalities by 2.51 times. According to a study on the severity of pedestrian injuries, the occurrence of collisions with pedestrians at intersections increases the probability of increasing pedestrian injuries (Haleem et al., 2015). These observations are consistent with the results of this study regarding the collision of buses and pedestrians at intersections.

Association rules for bus crashes on weekends (Cluster 3) are shown in Table 7, which satisfies the LIC criterion. The fatality rate of this cluster (5.2%) is more substantial than other clusters. According to Rule 1, crashes that occur on weekends and in darkness conditions increase the probability of fatalities by 5.9 times. According to previous studies, occurrence of crashes at night (Prato & Kaplan, 2014; Sam et al., 2018) and the time interval between midnight and dawn (Chu, 2014) have increased the severity of injuries in bus crashes. According to Rule 2, if the speed limit is increased in the darkness condition, the probability of fatality is 6.39 times more than the average. In several studies on the severity of bus crashes, there has been a direct relationship between the high speed limit and the severity of bus crashes (Nasri & Aghabayk, 2020; Prato & Kaplan, 2014) and that high speed limit increases the probability of bus crash occurrence (Chimba et al., 2010) and the severity of bus involved crashes (Kaplan & Prato, 2012).

According to Rule 3, bus accidents on highways increase the chances of fatality by 1.4 times. Although the presence of pedestrians on highways are usually not permitted, according to Rule 4, if such a situation occurs, the probability of fatalities increases 15.35 times. High probability of fatalities due to bus collisions with pedestrians on highways can be due to the vulnerability of pedestrians (Damsere-Derry et al., 2017) and the high speed of buses on highways.

According to Rule 5, female bus drivers increase the probability of fatalities by 3.38 times and also, according to Rule 6, if a bus with a female driver collides in the absence of special traffic control, the probability of fatalities increases by 4.36 times. In addition, rule 7 shows that the presence of heavy vehicles increases the probability of fatalities by 4.80 times. According to previous

Table 7
List of interesting rules identified for data in cluster 3 (Weekend).

LIC	Lift	C%	S%	Consequent	Antecedent	ID
n.a.	5.90	30.0	1.3	Fatality	Light condition = "Dark No street lights"	1
1.08	6.39	33.0	1.3	Fatality	Light condition = "Dark No street lights" & Speed limit = "Over50"	2
n.a.	2.18	11.3	1.6	Fatality	Road type = "Highway"	3
7.04	15.35	80.0	1.4	Fatality	Road type = "Highway" & DCA = "Struck Pedestrian"	4
n.a.	3.38	17.6	1.9	Fatality	Driver gender = "Female"	5
1.28	4.36	22.7	1.6	Fatality	Driver gender = "Female" & Traffic control = "No control"	6
1.10	4.80	25.0	1.3	Fatality	Driver gender = "Female" & Traffic control = "No control" & Heavy vehicle presence= "Heavy vehicle exist"	7
n.a.	1.23	6.5	4.8	Fatality	Speed limit = "Over50"	8
3.39	4.17	21.7	1.6	Fatality	Speed limit = "Over50" & Driver gender = "Female"	9
1.27	5.32	27.7	1.6	Fatality	Speed limit = "Over50" & Driver gender = "Female" & No of vehicles = "More than 1 vehicle"	10

studies, the gender of the bus driver is affected by the severity of the crashes (Feng et al., 2016) and that the female bus driver increases the severity of the injury (Kaplan & Prato, 2012).

According to Rule 8, if the speed limit is above 50 km/h, the probability of fatalities is 1.23 times more than the average. These results are consistent with the results of the direct effect of speed limit on the severity of bus crash injuries (Chimba et al., 2010). Additionally, according to Rule 9, if the speed limit is above 50 km/h and the bus driver is female, the probability of fatalities is 4.17 times more than the average. In addition, according to Rule 10, if a female bus driver collides with more than one vehicle in conditions of high-speed limit, the probability of fatalities will be 5.32 times more than average. The impact of the number of vehicles involved in the severity of bus crashes has been investigated and emphasized (Feng et al., 2016).

6. Conclusion

In this study, a combination of two data mining methods has been used to investigate the relationship between the bus involved crash severity and chains of the effective factors (environment, humans, and vehicles) that lead to a fatality in bus involved crashes. In the first stage, two-step clustering of crash data was separated, and in the second stage, using association rule discovery, the valid rules leading to an increase in the probability of fatalities in bus involved crashes were extracted. Following are the key observations from the model:

- Based on the clustering results, the factors that separate crash records into homogeneous groups with the different patterns are: the type of bus collision (collision with a motor vehicle or other types of collisions), day of the week.
- Bus crashes with motor vehicles on weekdays have caused a fatality rate of 1.6%, which is lower than the average rate (2.7%). According to the results, bus crashes with cars during weekdays, at roads with high-speed limit and collisions with more than one vehicle increase the probability of fatalities.
- All types of bus collisions except collisions with motor vehicles (Cluster 2), have resulted in a fatality rate of 3.2%, which is almost twice the fatality rate of bus collisions with motor vehicles (Cluster 1). The old age of the bus driver, the presence of vulnerable road users, and the collision of the bus with pedestrians at signalized intersections increase the probability of fatalities in Cluster 2.
- Cluster 3 includes all bus involved crashes on weekends with a high rate of fatality (5.2%), which indicates the difference in its pattern with bus crashes on weekdays. Darkness increases the probability of fatality on weekends, especially in areas with higher speed limits. Direct bus collisions with pedestrians on highways greatly increase the probability of fatalities. In addition, the female gender of the bus driver, the lack of special traffic controls, the presence of heavy vehicles, and collisions with multi vehicles are other factors affecting the increase in bus involved crash fatalities.

The results of this study show that more engineering and education efforts and coordination should be made to reduce the severity of bus involved crashes. Due to the special conditions of the bus in terms of larger size and more limited driver's vision than other vehicles, it is recommended to have awareness and safety training programs on rules, rights, and duties of other road users, including observing safe longitudinal and transverse distances with the buses. According to the results of this study, there is a high probability of pedestrian fatalities with bus crashes at traffic intersections. This may be due to distracted behaviors of the pedestrian at the intersection (Shiwakoti et al., 2019). Also, safeguarding bus maneuvers and bus turnings at urban intersections are part of the solution for pedestrian safety. The use of reflective accessories or brightly colored clothes (Hagel et al., 2007; Kwan & Mapstone, 2006) for vulnerable users like pedestrians can reduce the probability of fatality. Studies have shown that low-cost road safety education campaigns for pedestrians can reduce distracted and jaywalking behavior at intersections (Bungum et al., 2005; Shiwakoti et al., 2019). Illuminated in-ground Light Emitting Diodes (LEDs) embedded in pathways are also likely to be effective at attracting the attention of distracted pedestrians, as found in a recent study in Australia (Larue et al., 2020). If social rules or enforcement measures are consistently applied, it may also help in influencing road user's behavior during weekend entertainment activities. In addition, the installation of pedestrian detection systems on buses may help the driver to reduce collisions with pedestrians. Also, assigning bus routes with the minimal collision with other users in low-speed areas to old and female bus drivers can be effective. However, in a similar study of Melbourne city bus crashes, it is recommended to assign routes comprising mainly divided roads as well as newer and shorter buses to less experienced drivers (Goh et al., 2014a). Also, separating bus routes and setting special lanes in Melbourne city reduced fatal and serious injury crashes (Goh et al., 2013) and collisions (Goh et al., 2014b). Thus, bus priority especially in crowded areas where the movement of vulnerable users is higher can be a cost-effective way to reduce fatalities and crash counts (Goh et al., 2014c). Since in this study, bus crashes in the darkness and lack of light are associated with increased fatality risk, the necessary action can include improving the lighting conditions along the bus routes. Higher speed limits increase the probability of fatalities on direct routes. Installation of speed bumps and rumble strips along the bus route can improve safety.

The results of this study provide an insight into the pattern of crashes and the severity of bus involved crashes, which can be useful for managers of bus companies and road authorities to develop appropriate strategic measures and actions to reduce fatalities. This study used the database of a single state in Australia. In future, similar studies can be carried out in different regions to increase confidence in the insights into the pattern and severity of bus involved crashes. It is also suggested that human factors that may affect the occurrence of crashes such as the health status of bus drivers and medical conditions to be considered in future studies.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare that they have no financial and personal relationships with other people or organizations that could inappropriately influence (bias) their work.

References

af Wåhlberg, A. E. (2004). The stability of driver acceleration behavior, and a replication of its relation to bus accidents. *Accident Analysis & Prevention, 36*, 83–92.

Agrawal, R., & Srikant, R., 1994. Fast algorithms for mining associationrules. Paper presented at the Proceedings of the 20th International Conference on Very Large Data Bases, VLDB, Santiago de Chile, Chile.

Albertsson, P., & Falkmer, T. (2005). Is there a pattern in European bus and coach incidents? A literature analysis with special focus on injury causation and injury mechanisms. *Accident Analysis & Prevention, 37*, 225–233.

Alikhani, M., Nedaie, A., & Ahmadvand, A. (2013). Presentation of clustering-classification heuristic method for improvement accuracy in classification of severity of road accidents in Iran. *Safety Science, 60*, 142–150.
Barua, U., & Tay, R. (2010). Severity of urban transit bus crashes in Bangladesh. *Journal of Advanced Transportation, 44*, 34–41.
Besharati, M. M., & Tavakoli Kashani, A. (2018). Which set of factors contribute to increase the likelihood of pedestrian fatality in road crashes? *International Journal of Injury Control and Safety Promotion, 25*, 247–256.
BITRE (2019). *Australian Road Deaths Database*. Bureau of Infrastructure, Transport and Regional Economics. https://www.bitre.gov.au.
Brenac, T., & Clabaux, N. (2005). The indirect involvement of buses in traffic accident processes. *Safety Science, 43*, 835–843.
Bungum, T. J., Day, C., & Henry, L. J. (2005). The association of distraction and caution displayed by pedestrians at a lighted crosswalk. *Journal of Community Health, 30*, 269–279.
Chimba, D., Sando, T., & Kwigizile, V. (2010). Effect of bus size and operation to crash occurrences. *Accident Analysis & Prevention, 42*, 2063–2067.
Chiu, T., Fang, D., Chen, J., Wang, Y., & Jeris, C. (2001). A robust and scalable clustering algorithm for mixed type attributes in large database environment. In *Proceedings of the seventh ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 263–268).
Chu, H.-C. (2014). Assessing factors causing severe injuries in crashes of high-deck buses in long-distance driving on freeways. *Accident Analysis & Prevention, 62*, 130–136.
Clabaux, N., Fournier, J.-Y., & Michel, J.-E. (2014). Powered two-wheeler drivers' crash risk associated with the use of bus lanes. *Accident Analysis & Prevention, 71*, 306–310.
Damsere-Derry, J., Palk, G., & King, M. (2017). Road accident fatality risks for "vulnerable" versus "protected" road users in northern Ghana. *Traffic Injury Prevention, 18*, 736–743.
Das, S., Dutta, A., Avelar, R., Dixon, K., Sun, X., & Jalayer, M. (2019). Supervised association rules mining on pedestrian crashes in urban areas: Identifying patterns for appropriate countermeasures. *International Journal of Urban Sciences, 23*, 30–48.
Feng, S., Li, Z., Ci, Y., & Zhang, G. (2016). Risk factors affecting fatal bus accident severity: Their impact on different types of bus drivers. *Accident Analysis & Prevention, 86*, 29–39.
Goh, K., Currie, G., Sarvi, M., & Logan, D. (2014). Factors affecting the probability of bus drivers being at-fault in bus-involved accidents. *Accident Analysis & Prevention, 66*, 20–26.
Goh, K. C., Currie, G., Sarvi, M., & Logan, D. (2014). Experimental microsimulation modeling of road safety impacts of bus priority. *Transportation Research Record, 2402*, 9–18.
Goh, K. C. K., Currie, G., Sarvi, M., & Logan, D. (2013). Road safety benefits from bus priority: An empirical study. *Transportation Research Record, 2352*, 41–49.
Goh, K. C. K., Currie, G., Sarvi, M., & Logan, D. (2014). Bus accident analysis of routes with/without bus priority. *Accident Analysis & Prevention, 65*, 18–27.
Hagel, B. E., Lamy, A., Rizkallah, J. W., Belton, K. L., Jhangri, G. S., Cherry, N., & Rowe, B. H. (2007). The prevalence and reliability of visibility aid and other risk factor data for uninjured cyclists and pedestrians in Edmonton, Alberta, Canada. *Accident Analysis & Prevention, 39*, 284–289.
Haleem, K., Alluri, P., & Gan, A. (2015). Analyzing pedestrian crash injury severity at signalized and non-signalized locations. *Accident Analysis & Prevention, 81*, 14–23.
Hamid, I. R. A., & Abawajy, J. H. (2014). An approach for profiling phishing activities. *Computers & Security, 45*, 27–41.
Jovanis, P. P., Schofer, J. L., Prevedouros, P., & Tsunokawa, K., 1991. Analysis of bus transit accidents: Empirical, methodological, and policy issues.
Kaplan, S., & Prato, C. G. (2012). Risk factors associated with bus accident severity in the United States: A generalized ordered logit model. *Journal of Safety Research, 43*, 171–180.
Kashani, A. T., & Besharati, M. M. (2017). Fatality rate of pedestrians and fatal crash involvement rate of drivers in pedestrian crashes: A case study of Iran. *International Journal of Injury Control and Safety Promotion, 24*, 222–231.
Kashani, A. T., Rabieyan, R., & Besharati, M. M. (2014). A data mining approach to investigate the factors influencing the crash severity of motorcycle pillion passengers. *Journal of Safety Research, 51*, 93–98.
Kumar, S., & Toshniwal, D. (2016). A data mining approach to characterize road accident locations. *Journal of Modern Transportation, 24*, 62–72.
Kwan, I., & Mapstone, J. (2006). Interventions for increasing pedestrian and cyclist visibility for the prevention of death and injuries. *Cochrane Database of Systematic Reviews*.
Larue, G. S., Watling, C. N., Black, A. A., Wood, J. M., & Khakzar, M. (2020). Pedestrians distracted by their smartphone: Are in-ground flashing lights catching their attention? A laboratory study. *Accident Analysis & Prevention, 134* 105346.
Li, Y., & Fan, W. D. (2019). Modelling severity of pedestrian-injury in pedestrian-vehicle crashes with latent class clustering and partial proportional odds model: A case study of North Carolina. *Accident Analysis & Prevention, 131*, 284–296.
Li, Y., Lan, F., & Chen, J. (2012). Numerical study on the influence of superstructure configuration on coach rollover resistance performance. *International Journal of Crashworthiness, 17*, 637–654.
Li, Y., Su, G., Zhang, X., Zhang, S., & Yuan, H. (2015). Analysis of school bus accidents in China. *Natural Hazards, 79*, 723–734.
López, G., Abellán, J., Montella, A., & de Oña, J. (2014). Patterns of single-vehicle crashes on two-lane rural highways in granada province, Spain: In-depth analysis through decision rules. *Transportation Research Record, 2432*, 133–141.
Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice, 44*, 291–305.
Martínez, L., Aparicio, F., García, A., Páez, J., & Ferichola, G. (2003). Improving occupant safety in coach rollover. *International Journal of Crashworthiness, 8*, 121–132.
Mohamed, M. G., Saunier, N., Miranda-Moreno, L. F., & Ukkusuri, S. V. (2013). A clustering regression approach: A comprehensive injury severity analysis of pedestrian–vehicle crashes in New York, US and Montreal, Canada. *Safety Science, 54*, 27–37.
Montella, A. (2011). Identifying crash contributory factors at urban roundabouts and using association rules to explore their relationships to different crash types. *Accident Analysis & Prevention, 43*, 1451–1463.
Montella, A., Aria, M., D'Ambrosio, A., & Mauriello, F. (2011). Data-mining techniques for exploratory analysis of pedestrian crashes. *Transportation research record, 2237*, 107–116.
Montella, A., Aria, M., D'Ambrosio, A., & Mauriello, F. (2012). Analysis of powered two-wheeler crashes in Italy by classification trees and rules discovery. *Accident Analysis & Prevention, 49*, 58–72.
Montella, A., de Oña, R., Mauriello, F., Riccardi, M. R., & Silvestro, G. (2020). A data mining approach to investigate patterns of powered two-wheeler crashes in Spain. *Accident Analysis & Prevention, 134* 105251.
Nasri, M., & Aghabayk, K. (2020). Assessing risk factors associated with urban transit bus involved accident severity: A case study of a Middle East country. *International Journal of Crashworthiness*, 1–11.
Pande, A., & Abdel-Aty, M. (2009). Market basket analysis of crash data from large jurisdictions and its potential as a decision support tool. *Safety Science, 47*, 145–154.
Prato, C. G., & Kaplan, S. (2014). Bus accident severity and passenger injury: Evidence from Denmark. *European Transport Research Review, 6*, 17–30.
Sam, E. F., Daniels, S., Brijs, K., Brijs, T., & Wets, G. (2018). Modelling public bus/minibus transport accident severity in Ghana. *Accident Analysis & Prevention, 119*, 114–121.
Schepens, P. J., Pauwels, A., Van Damme, P., Musuku, A., Beaucourt, L., & Selala, M. I. (1998). Drugs of abuse and alcohol in weekend drivers involved in car crashes in Belgium. *Annals of Emergency Medicine, 31*, 633–637.
Seyedi, M., Jung, S., Dolzyk, G., & Wekezer, J. (2019). Experimental assessment of vehicle performance and injury risk for cutaway buses using tilt table and modified dolly rollover tests. *Accident Analysis & Prevention, 132* 105287.
Shiwakoti, N., Tay, R., & Stasinopoulos, P. (2019). Development, testing, and evaluation of road safety poster to reduce jaywalking behavior at intersections. *Cognition, Technology & Work*, 1–9.
Strathman, J. G., Wachana, P., & Callas, S. (2010). Analysis of bus collision and non-collision incidents using transit ITS and other archived operations data. *Journal of safety research, 41*, 137–144.
Sun, M., Sun, X., & Shan, D. (2019). Pedestrian crash analysis with latent class clustering method. *Accident Analysis & Prevention, 124*, 50–57.
Svetina, M. (2016). The reaction times of drivers aged 20 to 80 during a divided attention driving. *Traffic Injury Prevention, 17*, 810–814.
Tse, L. Y., Hung, W. T., & Sumalee, A. (2014). Bus lane safety implications: A case study in Hong Kong. *Transportmetrica A: Transport Science, 10*, 140–159.
Tseng, C.-M. (2012). Social-demographics, driving experience and yearly driving distance in relation to a tour bus driver's at-fault accident risk. *Tourism Management, 33*, 910–915.
Vajari, M. A., Aghabayk, K., Sadeghian, M., & Shiwakoti, N. (2020). A multinomial logit model of motorcycle crash severity at Australian intersections. *Journal of Safety Research*.
VicRoads, C.s., 2019. https://www.vicroads.vic.gov.au/safety-and-road-rules/safety-statistics/crash-statistics.
Wei, C., Becic, E., Edwards, C., Graving, J., & Manser, M. (2014). Task analysis of transit bus drivers' left-turn maneuver: Potential countermeasures for the reduction of collisions with pedestrians. *Safety Science, 68*, 81–88.
Weng, J., Zhu, J.-Z., Yan, X., & Liu, Z. (2016). Investigation of work zone crash casualty patterns using association rules. *Accident Analysis & Prevention, 92*, 43–52.
WHO, 2019. World health statistics 2019: monitoring health for the SDGs, sustainable development goals.
Xiaoyun, Z., Dongming, Z., Xiaobo, Y., & Xinyi, H. (2019). Improvement on design and manufacture of bus based on traffic accident prevention. *International Journal of Crashworthiness, 24*, 523–532.
Yoon, S., Kho, S.-Y., & Kim, D.-K. (2017). Effect of regional characteristics on injury severity in local bus crashes: Use of hierarchical ordered model. *Transportation Research Record, 2647*, 1–8.
Yu, R., & Abdel-Aty, M. (2013). Investigating the different characteristics of weekday and weekend crashes. *Journal of Safety Research, 46*, 91–97.
Zegeer, C.V., Huang, H., Hummer, J., Stutts, J., & Rodgman, E., 1993. Characteristics and solutions related to bus transit accidents.
Zhang, T., Ramakrishnan, R., & Livny, M. (1996). BIRCH: An efficient data clustering method for very large databases. *ACM Sigmod Record, 25*, 103–114.

Seyed Alireza Samerei is a PhD student in the Department of Civil Engineering, College of Engineering at the University of Tehran. His research interests are traffic safety analysis, data mining techniques and human factors in traffic engineering.

Kayvan Aghabayk is an Assistant Professor of School of Civil Engineering at the University of Tehran. His research interests are road safety engineering, vehicle collision research and traffic engineering. Prior to joining academia, Dr. Aghabayk worked for a range of organizations.

Amin Mohammadi is currently a PhD student in Civil Engineering at the University of Tehran. He is also a lecturer of Mianeh Technical and Engineering Faculty at the University of Tabriz. He is engaged in the research fields such as traffic safety, transportation planning, transportation systems and railway transportation.

Nirajan Shiwakoti is a transportation engineer with over 12 years of academic and professional experience. He is a Senior Lecturer at School of Engineering, Royal Melbourne Institute of Technology (RMIT) University. His research interests are multi-modal transport systems, logistics and supply chain, intelligent transport systems and traffic Safety.

Journal of Safety Research 76 (2021) 1–8

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Does seeing it make a difference? The self-reported deterrent impact of random breath testing

James Freeman *, Alexander Parkes, Verity Truelove, Naomi Lewis, Jeremy D. Davey

Road Safety Research Collaboration, University of the Sunshine Coast, Sippy Downs, 4556 Queensland, Australia

ARTICLE INFO

Article history:
Received 13 May 2020
Received in revised form 18 June 2020
Accepted 24 September 2020
Available online 3 November 2020

Keywords:
Drink driving
Random breath testing
Exposure

ABSTRACT

Introduction: Random Breath Testing (RBT) remains a primary method to both deter and apprehend drink drivers, yet a large proportion of road fatalities continue to be caused by the offense. Outstanding questions remain regarding how much exposure to RBT operations is needed to influence deterrence-based perceptions and subsequent offending. *Method:* Given this, licensed motorists ($N = 961$) in Queensland were recruited to complete a questionnaire either in the community ($N = 741$) or on the side of the road after just being breath tested ($N = 243$). Survey items measured different types of exposure to RBT operations (e.g., "seen" vs. "being tested") and subsequent perceptions of apprehension as well as self-reported drink driving behaviors. *Results:* The key findings that emerged were: motorists were regularly exposed to RBT operations (both viewing and being tested), such exposure was not significantly correlated with perceptions of apprehension certainty, and a sizable proportion reported engaging in drink driving behaviors (e.g., approx. 25%), although roadside participants naturally reported a lower percentage of offending behaviors. Importantly, it was revealed that current "observations" of RBT was sufficient, but not actual levels of active testing (which needed to be doubled). Nevertheless, higher levels of exposure to RBT operations was found to be predictive of a lack of intention to drink and drive again in the future. *Conclusions:* This paper suggests that mere exposure to enforcement may not create the intended rule compliance, and that the frequency of exposure is also essential for the roadside.

1. Introduction

Drink driving is a pervasive problem in many motorized countries, with clear links to increased crash risk and personal injury. Despite sustained enforcement efforts to reduce this behavior, in Australia 25% of road crash fatalities and 9.9% of road crash injuries are a result of drink driving (Transport and Main Roads, 2018). While a sizeable body of research has focused on the deterrent impact of drivers' perceptions toward police enforcement (e.g., Freeman & Watson, 2006; Grosvenor, Toomey, & Wagenaar, 1999; Homel, 1988; Szogi et al., 2017), there has been limited consideration on the extent to which drivers have actually been exposed to enforcement and the subsequent deterrent effect such exposure has on promoting rule compliance. This can be considered a significant omission not least because: (a) police enforcement (e.g., random breath testing) remains a primary method to deter motorists from drink driving; (b) there is a need to maximize ever restricting police resources; and (c) outstanding questions remain regarding whether merely observing enforcement (rather than actually being tested) can create a deterrent effect. Consequently, this study focuses on the impact of exposure to legal enforcement, such as Random Breath Testing (RBT), on self-reported engagement in drink driving behavior(s).

The most dominant enforcement method for deterring drink driving remains RBT (primarily through general deterrence), and the mechanism has repeatedly been demonstrated to have a positive effect on road safety in Australia. For example, earlier research demonstrated that after the introduction of RBT, alcohol related road crash fatalities decreased by 18% (Watson, 1994). Similarly, increasing the number of motorists tested (in Queensland) reduced both alcohol-related crashes and drink driving offense rates (Watson et al., 2005). Similar results have been found in the United States, where instead of RBT, there are sobriety checkpoints that target drivers who are perceived as being impaired by alcohol (instead of selecting drivers at random; Bergen et al., 2014). It was found that there were 18% less cases of drink driving in states that have implemented sobriety checkpoints in comparison to states that have not implemented these checkpoints (Lenk, Nelson, Toomey, Jones-Webb, & Erickson, 2016).

* Corresponding author at: University of the Sunshine Coast, Sippy Downs Campus, Maroochydore DC, QLD 4558, Australia.

E-mail address: jfreema1@usc.edu.au (J. Freeman).

https://doi.org/10.1016/j.jsr.2020.09.013
0022-4375/

However, conducting RBTs remain costly both in regards to time and resources, and a sizable proportion of fatalities are still attributed to drink driving despite the intensive application of RBT operations. For example, in 2018, while 7,921 RBTs were conducted to every 10,000 licensed drivers in Queensland (BITRE Bureau of Infrastructure Transport and Regional Economics, 2019), 17.6% of Queensland driver fatalities involved a drink driver (Department of Transport and Main Roads, 2018). Given that the previously cited target ratio of 1:1 RBT conducted to licensed driver (Ferris et al., 2013) is not sustainable into the foreseeable future, it is timely to consider what type and level of RBT intensity is needed to provide (and maintain) a deterrent impact on motorists.

Interestingly, and despite the widespread reliance on RBT testing (in Australia), the optimum level of application has not currently been established (Ferris et al., 2013; Lenk et al., 2016), although as noted above, the 1:1 ratio has been utilized as a broad goal. That is, there is no benchmark for exposure to RBT operations, and questions remain if "observing" an RBT operation enacts the same deterrent effect as actually being breath tested. If this can be proven to be the case, then operating highly visible RBT operations may prove to be an excellent proxy method for maintaining a deterrent effect. At best, a preliminary study by Morrison, Ferris, Wiebe, Peek-Asa, and Branas (2019) that analyzed crash rates with roadside sobriety checkpoints found that the effect of sobriety checkpoints was maintained for one week. Additionally, a United States study reported a 41% reduction in drink driving was identified when the rate of sobriety checkpoints was increased to monthly, in comparison to states that did not conduct sobriety checkpoints (Lenk et al., 2016). Conversely, one study that examined the impact of exposure to RBT on drink driving engagement, including observing RBT as well as actually being tested by an RBT, found that these variables did not predict self-reported drink driving events (Watson & Freeman, 2007). However, the sample's actual exposure to RBT was low.

Taken together, little is known about the "perceptual" impact of exposure to RBT operations, although a complementary Queensland-based study on exposure to speed cameras failed to find a strong link with speed limit compliance (Freeman, Kaye, Truelove, & Davey, 2017). Nevertheless, and stemming from models of learning, it may be suggested that drivers still need to be exposed to high levels of enforcement in order to be effectively deterred from drink driving. That is, learning from experience and exposure are crucial for humans (Meeter, Shohamy, & Myers, 2009) and the principles of operant and Pavlovian conditioning have been extensively validated in the scientific literature (Britton, Lissek, Grillon, Norcross, & Pine, 2011). At the very least, the importance of the relationship between stimulus and response have been well documented, particularly in regards to frequent exposure (Recio, Iliescu, & de Brugada, 2019). In regards to the question of "being tested" versus merely "observing RBT testing," more may be learned from behavioral psychology that has demonstrated exposure to two similar stimuli may reduce generalizations between them (Sanjuan & Nelson, 2019). On the other hand, length of exposure to a stimulus also remains important (Recio et al., 2019), which suggests actually being breath tested may produce superior deterrent effects (rather than just driving past an RBT operation). Such issues will be considered in this exploratory study.

Within the domain of road safety, exposure to enforcement may (most crucially) be linked to increasing perceptions of certainty of apprehension, which has long been hypothesized to be the most important of the three classical deterrent forces (Von Hirsch & Colloquium, 1999). This construct refers to the idea that if an individual believes there is a high chance of being caught and punished for committing a crime, they will subsequently be deterred from committing that offense (Beccaria, 1764/2007; Bentham, 1780/1970). Put simply, if a driver has limited exposure to RBT, it may be suggested that they consequently have low perceptions of the certainty of being caught, which dilutes deterrent forces and promotes drink driving. Conversely, if a driver has higher exposure to RBT (resulting in a higher perception of the certainty of being caught), they may be more deterred from drink driving. On the one hand, previous research surrounding perceptual certainty of apprehension and drink driving found that certainty of apprehension was a significant negative predictor of drink driving (Freeman & Watson, 2006, 2009). Alternatively, other studies have found this variable not to be a direct predictor of drink driving (Baum, 1999; Homel, 1988; Watson & Freeman, 2007). Such discrepancies may yet be explained by moderating forces such as actual exposure to enforcement.

1.1. The current study

Given the above, outstanding questions remain regarding the perceptual deterrent impact of exposure to RBT on subsequent offending behaviors. This study had the following research questions:

1. What is the deterrent impact of RBT exposure on perceptual certainty of apprehension for drink driving?
2. Is there a differential effect between being tested versus observing testing?
3. How does exposure to enforcement impact upon self-reported drink driving behaviors and what other factors predict the offense?

A secondary (exploratory) research question focused on examining whether perceptual differences and self-reported offending behaviors differed regarding the period of time since most recent exposure to RBT. That is, are differences identifiable between individuals who had recently been tested (e.g., straight after an RBT) versus in the community (e.g. recency effect).

2. Method

2.1. Participants

A total of 961 participants took part in the study as part of a larger state-wide deterrence-based research project. The sample was comprised of 714 community members who were recruited either online or face to face in public places (e.g., university campuses, libraries, and shopping centers). The remaining 243 participants were drivers of motor vehicles recruited at the roadside immediately following a random roadside breath test (outlined in Table 1). No significant sociodemographic differences were found for age, years driving, or the distribution of males and females across the roadside and community groups. Weekly driving hours were also fairly consistent between the two groups.

2.2. Measures

2.2.1. Participant characteristics

Participants were asked to provide demographic data pertaining to age, gender, years of licensure, type of license held, weekly driving hours, and previous apprehension for drink driving offenses.

2.2.2. Drink driving behavior

Consistent with the need to expand the methodological operationalization of drink driving events (Freeman et al., 2020), a multi-

Table 1
Participant characteristics for the community and roadside samples.

	Combined sample		Community		Roadside	
	n [M]	*% [SD]*	*n [M]*	*% [SD]*	*n [M]*	*% [SD]*
Total	961	100%	718	74.71%	243	25.29%
Age	[39.55]	[16.04]	[39.61]	[16.33]	[39.34]	[15.17]
Years driving	[20.41]	[16.03]	[20.26]	[16.35]	[20.85]	[15.08]
Male	502	52.24%	371	51.67%	132	54.32%
Female	453	47.14%	346	48.20%	110	45.27%
Weekly driving hours						
<5	196	20.40%	153	21.30%	43	17.70%
6–10	305	31.74%	235	32.70%	70	28.81%
11–20	234	24.35%	162	22.60%	72	29.63%
21–30	95	9.89%	74	10.30%	21	8.64%
>30	125	13.01%	91	12.70%	34	13.99%

item approach was utilized to examine the frequency of engagement in self-reported drink driving events that were: (a) a possible drink driving episode, e.g., *"How often do you drink and drive when you <u>think</u> you may have been over the legal blood alcohol limit?*" and (b) an acknowledged episode, e.g., *"How often do you drink and drive when you <u>know</u> you have a blood alcohol level above the legal limit?*" An additional proxy measure of drink driving was also included (*"How often have you attempted to avoid Random Breath Testing sites when you see them?"),* as well as a measure of future intentions to drink drive with one item ("*It's likely that I will drive above the legal blood alcohol limit in the future*"). Item responses were collected on a 7-point Likert scale ranging from 1 (*never*) to 7 (*always*) for both possible and acknowledged drink driving and evading RBT sites and 1 (*strongly disagree*) to 2 (*strongly agree*) for future intentions to drink and drive. Higher scores for all drink driving variables indicated more frequent self-reported drink driving behaviors/intentions.

2.2.3. RBT exposure

Exposure to RBT operations was quantified for both seeing an RBT site; *"In the past 12 months, how often have you SEEN the police conducting random breath testing (even if you weren't tested)*" and being breath tested *"In the past 12 months how many times have you been stopped and breath tested by police?*" To estimate participant's own perceptions of the deterrent effect of RBT exposure, two subjective ratings were included, "*How often would you need to see an RBT over a 12-month period to make you avoid drink driving*" and "*How often would you need to be breath tested at an RBT site over a 12 month period to make you avoid drink driving.*"

2.2.4. Apprehension certainty

Consistent with previous research (Freeman et al., 2020), an objective measure of certainty of apprehension for drink driving *"The chances for getting caught for drink driving are high,"* was utilized with response items on a 7-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*).

2.2.5. Alcohol consumption and drinking behavior

Participant's behavior relating to alcohol consumption was classified as either non-risky or risky drinking according to the 4-item short form of the Alcohol Use Disorder Identification Test (AUDIT-4; Gual, Segura, Contel, Heather, & Colom, 2002). An item example is *"How often do you have six of more drinks on one occasion?*" items are scored from 0 – 4 (e.g., 0 = "*never,*" 1 = "*less than monthly,*" 2 = "*monthly,*" 3 = "*weekly,*" and 4 = "*daily or almost daily*"). Scores for each item are added to give a total score ranging from 0 to 16. A score > 7 for males and > 5 for females indicates risky drinking behavior.

2.3. Procedure

For both the community and roadside populations, a representative sample of Queensland motorists was sought by conducting surveys across a balanced selection of cities (e.g., Logan, Rocklea, Gold Coast, Townsville, and Ipswich) and regional councils (e.g., Toowoomba, Gympie, Emerald [Central Highlands], and Rockhampton). A $20 Coles/Myer gift voucher was offered as reimbursement for participation. Community sample participants either completed a paper survey in-person at various community settings such as the Queensland University of Technology campus, shopping centers, and other public places. Roadside participants were offered the opportunity to participate in the study after being randomly breath tested in various locations throughout Queensland. That is, volunteering participants completed the surveys (in their vehicles) approximately 20 meters further down the roads (from the operational RBT site). No identifying information was recorded to ensure the anonymity and confidentiality of responses. The collected survey data was entered into the IBM SPSS statistics (version 24) program for analysis.

3. Results

3.1. Self-reported drink driving

Self-reported drink driving behaviors were relatively uncommon (refer to Table 2). In regards to self-reported offending for the total combined sample, "acknowledged" drink driving (e.g., knowingly over the limit) 827 participants (86.1%) responded with the "never" category. The same proportion of respondents, $n = 827$, (86.1%) reported they had "never" attempted to evade RBT sites to avoid potential breath testing by police (e.g., a proxy measure of drink driving behavior) in the 12 months prior to the survey. Additionally, the largest proportion (e.g., $n = 749$, 77.9%) reported "never" driving while "thinking" they may be over the legal limit, which is consistent with previous research (Freeman, Szogi, Truelove, & Vingilis, et al., 2016).

To capture the sum total of reported drink driving behavior, a dichotomized variable was constructed by combining the three items (e.g., possible, acknowledged, and proxy) such that respondents who answered "never" on all three variables were designated as a "no" condition or not drink drivers, and all other responses were categorized as "yes" or, drink drivers. The overall frequency of participants who were consequently not drink drivers was 695 (72.3%).

However, and from a different perspective, it is noteworthy that more than 25% of the sample may have engaged in drink driving events, which reinforces the significance of the ongoing drink driving problem. This is further evidenced by an examination of inten-

Table 2
Self-reported drink driving behavior.

Construct	Survey Items	M	SD	p	Never 1	Rarely 2	Sometimes 3	Uncertain 4	Often 5	Nearly Always 6	Always 7
Possible Drink Driving	*How often do you drive when you think you may have been above the legal blood alcohol limit?*	1.31	0.72		77.9%	15.9%	4.9%	0%	0.9%	0%	0.3%
Community		1.36	0.79	<0.001	74.8%	18.1%	5.4%	0%	1.3%	0%	0.4%
Roadside		1.16	0.45		87.2%	9.5%	3.3%	0%	0%	0%	0%
Acknowledged Drink Driving	*How often do you drive when you know you have a blood alcohol level above the legal limit?*	1.21	0.63		86.1%	8.1%	5.0%	0%	0.5%	0%	0.2%
Community		1.26	0.70	<0.001	84.0%	8.6%	6.3%	0%	0.7%	0%	0.3%
Roadside		1.09	0.33		92.2%	6.6%	1.2%	0%	0%	0%	0%
Proxy Drink Driving Measure	*How often have you attempted to evade RBT sites when you see them?*	*1.27*	0.83		86.1%	5.7%	5.7%	0.6%	0.5%	0.4%	0.6%
Community		*1.33*	0.91	<0.001	83.4%	6.1%	7.4%	0.8%	0.7%	0.6%	0.7%
Roadside		*1.09*	0.47		93.8%	4.5%	0.8%	0%	0%	0%	0.4%

Note: p = significance of *t*-test for the null hypothesis that the mean score for the roadside sample is equal to that of the community.

tions to drink and drive again in the future, with only 76% of the combined sample responding with "strongly disagree" or "disagree" to such a likelihood (e.g., 24% were unwilling to confirm such an outcome). A smaller percentage of respondents reported having been apprehended (at any time) for drink driving in the past (n = 115, 12.0%). A corresponding chi square analysis revealed no significant association between drink driving apprehension and the two sampled cohorts.

A comparative analysis between the two groups (as outlined in Table 3) revealed statistically significant differences for self-reported drink driving behaviors, with roadside participants reporting the lowest frequency. This suggests that motorists may have (understandably and not surprisingly) engaged in a higher level of self-report bias when in close proximity to police personnel (e.g., roadside). However, there was no significant difference found for future intentions to drink drive between the roadside and community samples.

3.2. Perceptions of apprehension certainty by RBT exposure

In order to examine the deterrent impact of RBT operations, an initial baseline assessment of perceptions of apprehension certainty was undertaken. The results were consistent with previous research by Freeman et al. (2016) and demonstrated that the most common response to the statement "*The chances of getting caught for drink driving are high*" was "*agree*" (n = 257, 26.7%). Whereas the least common (n = 86, 8.90%) was both "*strongly disagree*" and "*neither agree nor disagree.*" However, overall, respondents tended to "*somewhat agree*" that a drink driving apprehension was certain (M = 4.57, SD = 1.85; the average score of four reflects a response of "*neither agree or disagree*" while five equals "somewhat agree"). Not surprisingly, a significant difference was identified between the community and roadside sample's reported levels of perceptual certainty. That is, an independent samples t-test revealed the average rating for certainty by the community sample (M = 4.37, SD = 0.07) was significantly lower than the average reported by the roadside sample (M = 5.15, SD = 0.11; $t(458.29) = -0.6.10$, $p < 0.001$). The higher rating for the roadside group highlights a potential recency effect of RBT exposure on the certainty of being apprehended for drink driving for this group.

In regards to RBT exposure, the vast majority of respondents (n = 898, 93.4%) reported being exposed to an RBT operation at least once in the previous 12 months (M = 4.18 SD = 2.71) with the most common frequency (mode) being two (n = 145, 12.1%). Among the population who had been exposed to an RBT, 292 (30.38% of the entire sample) had only seen an RBT (and not been tested) on average 2.8 times (SD = 2.49). In comparison 355 (36.9%) had not been tested and, of these, 63 participants (6.6% of the total sample) had not been exposed at all (seen or tested). As expected, being tested at an RBT site was less frequent than merely seeing an RBT. More specifically, n = 606 (63.1%) were tested at least once (M = 1.36, SD = 1.38) with responses ranging from once (n = 226, 23.5%) to four times (n = 114, 10.8%) and a mode of 1 (n = 226, 23.5%). It may be hypothesized that roadside participants reported a higher frequency of RBT exposure, not least because of the most recent event (e.g., exposure prior to completing the survey). A corresponding between groups (t-test) examination revealed that the mean number of times seeing an RBT for the community group (M = 4.1, SD = 2.9), was not significantly different than this type of exposure reported by roadside sample (M = 4.42, SD = 2.5; $t(959) = -1.55$ p = 0.122). In contrast and as expected, a similar investigation indicated that the community sample's average

Table 3
Dichotomised drink driving variables.

Drink Driver	Total		Community		Roadside		Chi-square	
	n	%	n	%	n	%	χ^2	p
Yes	266	27.7	224	31.2	42	17.3	17.56	<0.001
No	695	72.3	494	68.8	201	82.7		
Future Intentions								
Yes	226	23.5%	180	25.1%	46	18.9%	3.81	0.054
No	735	76.5%	538	74.9%	197	81.1%		

Note: Drink driver "*no*" is categorized by "*never*" on three drink driving variables (think, know and evade RBT); Future intentions "*no*" is categorized by responses "*strongly disagree*" and "*disagree*" on the single 7-point Likert item. Chi-square comparison is between the community and roadside samples.

Table 4
Actual exposure compared to estimated level needed for a deterrent effect.

RBT exposure (actual and *estimated)	Combined		Community		Roadside		p
	M	SD	M	SD	M	SD	
Actual Seen and not tested	2.94	2.64	2.9	2.7	3.1	2.4	0.603
Actual Seen and may have been tested	4.18	2.71	4.1	2.9	4.4	2.5	0.122
*Need to see an RBT	3.26	2.46	3.35	2.47	2.98	2.4	0.053
Actual Tested	1.36	1.38	1.28	1.38	1.61	1.34	<0.001
*Need to be Tested	2.39	2.26	2.45	2.3	2.21	2.15	0.150

Note: *self-reported estimates of exposure frequency required to act as a deterrent to drink driving; p = significance levels for 2-tailed independent samples t-test.

levels of exposure to an actual RBT test (M = 1.28, SD = 1.38) was significantly lower than reports from the roadside sample (M = 1.61, SD = 1.34; $t(427) = -3.26$, p = 0.001).

An additional exploratory analysis was undertaken (refer to Table 4) to examine how much RBT exposure was estimated by the sample to be needed to create a deterrent effect (measured by making a participant avoid drink driving). In response to the question '*how often would you need to see an RBT site over a 12-month period to make you avoid drink driving,'* responses ranged between 0 and 10 with a mean of 3.33 (SD = 2.48) and a mode of 6. Meanwhile, for the question '*how often would you need to be breath tested at an RBT site over a 12-month period to make you avoid drink driving,'* the responses also ranged between 0 and 10 with a mean of 2.46 (SD = 2.31) and a mode of 1 (responses for the roadside and community sample were not dissimilar and no statistical difference was found for mean ratings of needing to see or be tested in order to be deterred from drink driving).

It is noteworthy that in a majority of cases, these (estimated) mean scores are above the sample's actual level of exposure to RBTs (tested and seeing without being tested at all) in the past 12 months. Although, for those who saw an RBT and may also have been tested, the average reported frequency of seeing an RBT was higher than estimates (refer to Table 4). In terms of exposure compared to the estimated deterrent level, the results show that only 203 participants (21.12%) had been tested at or above the average estimated deterrent level (e.g., $\geq$1.36). In contrast, 511 participants (53.17%) had seen an RBT (who also may have been tested) at or above the average estimated deterrent level (e.g., $\geq$3.26). Whereas, for those who had not been tested at all, only 95 (9.89%) had seen an RBT at or above the frequency estimated to be a deterrent.

RBT exposure was further investigated to confirm whether increases in the number of weekly driving hours corresponded with higher levels of RBT exposure. As expected, visual observation of means showed higher frequencies of both seeing an RBT and being tested as weekly driving hours increased (see Fig. 1).

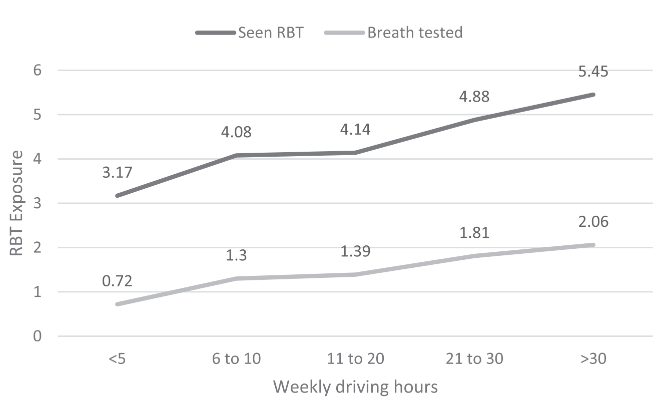


Fig. 1. Means plot demonstrating the observed linear relationship between weekly driving hours and RBT exposure; seen RBT category n = 319 does not include being breath tested n = 418.

3.3. Bivariate correlations

Pearson's r bivariate correlations between variables for the community group are presented in Table 5. Please note the roadside participants were excluded from further analysis due to the potential unreliability of self-reported drink driving behavior examined above. Of note, a positive and significant (albeit weak) relationship was found between RBT exposure (both observations and being tested) with certainty of apprehension. This relationship implies that more exposure to RBT operations increases perceptual certainty of apprehension. Although encouraging, certainty was not found to have a tangible impact on self-reported offending behaviors (see below). As expected, moderate significant positive correlations were found between past drink driving behavior and future intentions to drink drive. Not surprisingly, drinking behavior (measured on the AUDIT-4) was also moderately correlated with past drink driving behavior and weakly correlated with future intentions to drink drive. Being apprehended in the past was significantly negatively correlated with past drink driving (with a moderate effect) as well as future intentions to drink drive (with a weak effect), which suggests that: (a) some offenders may be immune/impervious to the threat (and application) of sanctions and/or (b) some offending behaviors may prove to be habitual in nature.

3.4. Multivariate analysis predicting apprehension certainty

To determine if levels of apprehension certainty could be predicted by key independent variables, an ordinal regression model was designed with predictors of age, drinking behavior, RBT exposure (seen and tested), weekly hours driving, gender, and past drink driving apprehension. The model produced an acceptable fit statistic and demonstrated goodness of fit, however only 5% of variance in apprehension certainty was accounted for (Nagalkerke R^2). The significant predictors in the model were: age, years of licensure, and seeing an RBT (see Table 6 for results).

3.5. Multivariate analyses predicting drink driving behavior

To analyze predictors of drink driving behavior (within the community sample), two separate logistic regression models were conducted (Table 7). The first model predicted past drink driving behavior dichotomized into a *"never"* condition (n = 495) or *"other than never"* condition (n = 224), thereby identifying individuals who acknowledge a past indiscretion. The second model predicted future intentions to drink drive with those who answered "*disagree*" or "*strongly disagree*" as the *"no"* condition (n = 538), and all other responses as a *"yes"* condition (n = 180). Each regression model had the same predictors of age, gender, hours driving, risky drinker [yes/no], apprehension certainty, breath tested frequency, seen RBT frequency, and drink driving apprehension. For the future intentions model, past drink driving behavior (dichotomized as previously described) was included to determine the predictive utility of past behavior on future intentions. A 2-stage hierarchical

Table 5
Bivariate correlations between variables for the community sample.

		1	2	3	4	5	6	7	8	9	10	11
1	Possible DD	1										
2	Acknowledged DD	0.82**	1									
3	Future Intentions to DD	0.31**	0.32**	1								
4	Evade RBT	0.60**	0.62**	0.25**	1							
5	DD apprehension	−0.29**	−0.35**	−0.13**	−0.23**	1						
6	Seen RBT	0.10**	0.09*	0.00	0.09*	−0.03	1					
7	Tested RBT	0.04	0.04	−0.03	0.07	−0.03	0.45**	1				
8	Certainty of apprehension	0.02	0.02	0.09*	0.00	−0.03	0.12**	0.10*	1			
9	Age	−0.13**	−0.12**	−0.1	−0.20**	−0.10*	0.00	0.01	0.09*	1		
10	Gender	−0.13**	−0.09*	−0.02	−0.04	0.12**	−0.03	−0.04	0.04	0.02	1	
11	AUDIT 4	0.39**	0.34**	0.13**	0.23**	−0.25**	0.10**	0.00	0.1	−0.11**	−0.26**	1
12	Road exposure	−0.04	0.03	0	0.02	−0.1	0.17**	0.21**	0.03	0.03	0.01	0.07

Note: ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).
Correlation between possible DD and certainty in the roadside sample was −0.18 ($p < 0.01$).
Community.

Table 6
Ordinal regression predicting certainty of apprehension.

	B	Wald	*P*	95% CI	
				Lower	Upper
Age	0.01	9.67	0.002**	0.01	0.02
Males	−0.22	2.44	0.119	−0.49	0.06
Females	0a	.	.	.	.
AUDIT 4	0.05	4.38	0.036*	0	0.09
Seen RBT	0.07	6.19	0.013*	0.01	0.12
Tested RBT	0.05	0.97	0.325	−0.05	0.16
Hours Driving = <5	−0.37	2.17	0.141	−0.86	0.12
Hours Driving = 6–10	−0.41	3.34	0.068	−0.86	0.03
Hours Driving = 11–20	−0.13	0.31	0.576	−0.6	0.33
Hours Driving = 21–30	−0.42	2.24	0.135	−0.97	0.13
Hours Driving = >30	0a	.	.	.	.
CaughtDD (N)	0.01	0	0.957	−0.41	0.44
CaughtDD (Y)	0a	.	.	.	.
Model fit (logit link function)			$\chi^2(10) = 2622.61$, $p < 0.001$		
Goodness of fit (Pearson Chi-Square)			$\chi^2(4208) = 4257.90$, $p = 0.291$		
Nagelkerke R^2			0.05		
Proportional odds assumption			$\chi^2(50) = 73.94$, $p = 0.015$		

Note: Results are for the community sample only; *B* = unstandardized coefficient representing the increase in the outcome variable scale for every one unit increase in the predictor; A comparative analysis for the roadside sample revealed that the only significant predictor to be: age ($B = 0.024$, $p = 0.003$).

regression model was used for each analysis, with demographic variables of age, gender, and hours driving in step 1 and all variables in the final model (fit statistics and goodness of fit are provided for step 2 of each model). The predictive value of each independent variable was measured as the odds ratio (OR), referring to the odds of being in the designated outcome category for each level increase of the predictor (holding all other variables constant). Both models produced acceptable fit statistics. The significant predictors of drink driving were younger motorists, gender (males), risky drinking, and a previous drink driving apprehension. In the future intentions model, only past drink driving behavior (drink driver yes category) and being a male were identified as significant predictors.

4. Discussion

This study aimed to examine the deterrent impact of reported exposure to RBTs on perceptions of certainty of apprehension and subsequently engagement in drink driving behavior. Three different ways of operationalizing drink driving were utilized to provide a more accurate and in-depth examination into the possible influence RBT exposure may have on behavior. The current study used a novel approach to recruitment to obtain two participant cohorts (e.g., general community members and those who had just been tested). This facilitated a comparative analysis of the certainty of apprehension post breath testing compared to a more general level of RBT exposure.

In regards to RBT exposure, and as expected, participants reported observing an RBT more than twice as often as being actually tested. However, and consistent with the Queensland Police Service's aim of testing every licensed driver once per year, the sample reported being tested annually (e.g., 1.36). However, a total of 63.1% of participants reported being breath tested during the previous 12 months, which is lower than the official published rate in 2018 (72.91%; BITRE, 2019). Interestingly, in regards to the exploratory analysis to identify the estimated level of exposure to create a strong deterrent effect, participants' current exposure levels (e.g., "seeing) were higher than what was proposed (e.g., $M = 4.18$ versus $M = 3.26$, respectively). In contrast, participants' actual level of being breath tested (e.g., 1.36) was approximately half the frequency that was stated to create a strong deterrent effect (e.g., 2.39 times). These results (to some degree) support the regression analysis finding that current levels of "seeing" an RBT had an effect on perceptual certainty, but not actual engagement in testing. Put differently, the frequency of exposure to being breath tested (e.g., once annually) may not be sufficient to counteract factors that promote drink driving (reviewed below). Based on these results, and considering limited police resources, it may be suggested that simply increasing the visibility of RBTs, instead of also increasing engagement in testing, may be sufficient to, at least

Table 7
Logistic regression analyses predicting drink driving outcome variables.

	Drink Driver Y/N					Future intentions to drink drive Y/N				
	B	*p*	OR	95% CI for OR		*B*	*P*	OR	95% CI for OR	
				Lower	Upper				Lower	Upper
Step 1										
Age	−0.02	<0.001***	0.98	0.97	0.99	−0.01	0.207	0.99	0.98	1
Gender (M)	−0.69	<0.001***	1.95	1.35	2.83	0.38	0.021*	1.46	1.06	2.02
Hours driving	−0.1	0.174	0.9	0.78	1.05	0.27	0.27	1.04	0.97	1.13
Step 2										
Age	−0.04	<0.001***	0.97	0.95	0.98	0	0.563	1	0.99	1.02
Gender (M)	0.61	0.001**	1.85	1.27	2.68	0.17	0.355	1.18	0.83	1.69
Hours driving	−0.01	0.844	0.99	0.94	1.06	0.06	0.343	1.06	0.94	1.19
Risky drinking (Y)	1.16	<0.001***	3.19	2.17	4.68	0.23	0.251	1.26	0.85	1.86
Apprehension certainty	0	0.994	1	0.91	1.1	0.07	0.156	1.07	0.97	1.18
Seen RBT	0.04	0.278	1.04	0.97	1.12	−0.02	0.55	1.02	0.95	1.1
Tested RBT	0.03	0.665	1.03	0.89	1.2	−0.14	0.062	0.87	0.76	1.01
DD apprehension (Y)	1.5	<0.001***	4.48	2.61	7.66	0.29	0.284	1.34	0.78	2.3
Drink Driver (Y)	–	–	–	–	–	1.67	<0.001***	5.32	3.57	7.92
Constant	−0.67	0.056	0.51			−1.87	0	0.15		
Model Chi-Square	$\chi^2(8) = 80.57$, $p < 0.001$					$\chi^2(9) = 108.42$, $p < 0.001$				
Hosmer Lemeshow	$\chi^2(8) = 5.25$, p = 0.731					$\chi^2(8) = 5.97$, $p = 0.651$				
Nagelkerke R^2	0.24					0.21				

Note: * α = 0.05; ** α = 0.01; *** α = 0.001; *C.I. = confidence interval; B* = unit change in the log of the odds of the outcome for every one unit change of the predictor; OR = odds ratio. Community sample only ($n = 718$).

partially, increase deterrence perceptions related to drink driving. This may be done by conducting RBTs at locations that have a high traffic density. At the very least, the assumption that greater exposure to the road would lead to greater exposure to RBT was supported (which supports the reliability of the results) and that those who were recently tested (e.g., roadside) also reported high levels of overall testing. The final analysis to focus on perceptual certainty revealed that being a younger motorist was predictive of lower levels of apprehension certainty, which is supportive of broader theories of psychological development that suggests younger individuals are more vulnerable to fail to recognize (and respond appropriately) to risk (Lebel & Beaulieu, 2011).

The complementary regression analyses to identify past predictors of past drink driving events revealed the usual suspects. That is, being younger (Freeman et al., 2020; Szogi et al., 2017), being male (Freeman et al., 2020; Szogi et al., 2017), consuming higher levels of alcohol (Freeman et al., 2020; Freeman & Watson, 2009), and previous drink driving convictions (Watson et al., 2017) were all predictive of past drink driving events. In contrast, the lack of identified predictors for future intentions to drink and drive may be more reflective of the array of forces (both personal and contextual) that can influence any possible drink driving event that naturally stems beyond the scope of the current study. Given that efforts to avoid RBT was a predictor of intentions to offend in the future, it may be suggested that past behavior remains a strong predictor of future behavior. As such, countermeasures that target recidivist offenders may be a useful allocation of resources to limit engagement in drink driving.

5. Concluding remarks

The study's limitations should be considered when interpreting the results that include the: (a) relatively small convenience sample, (b) heavy reliance on self-report data and Likert scales, and (c) the current sample may not be reflective of the broader Queensland motoring population. Despite such limitations, the current study contributes to a body of research that has failed to find strong links between enforcement and drink driving outcomes (Szogi et al., 2017; Watson et al., 2017) and is similar to a sister-study that failed to find a strong relationship between drivers' exposure to speed cameras and subsequent driving behaviors (Freeman et al., 2017). Such counterintuitive results may reinforce the array of personal and environmental factors that influence any one breach, not least, the habitual nature of the driving task that is often overlooked in the empirical literature. From a different perspective, it is noted that deterrence-based models are based on utility and rational decision making, which, respectfully, may not be a core construct of those who operate motor vehicles. Overall, based on the results from this study, it may be suggested that increasing visibility of RBTs, as well as targeted interventions aimed at recidivist drink drivers, may be beneficial to decrease drivers' engagement in drink driving.

6. Declarations of interest

None.

Acknowledgement

This research was funded by the Motor Accident Insurance Commission (MAIC).

References

Baum, S. (1999). Self-reported drink driving and deterrence. *Australian & New Zealand Journal of Criminology, 32*, 247–261.
Beccaria, C. (1764/2007). On crimes and punishments. In R. Davies & V. Cox (Trans.), R. Bellamy (Ed.). In: Cambridge, UK: Cambridge University Press. DOI.
Bentham, J. (1780/1970). An Introduction to the Principles of Morals and Legislation, eds. In J. H. Burns & H. L. Hart (Eds.), The collected works of Jeremy Bentham. University of London/The Athlone Press, London (Bellamy, R ed.).
Bergen, G., Pitan, A., Qu, S., Shults, R. A., Chattopadhyay, S. K., Elder, R. W., et al. (2014). Publicized sobriety checkpoint programs. *American Journal of Preventive Medicine, 46*, 529–539.
BITRE Bureau of Infrastructure Transport and Regional Economics (2019). *Random breath testing in 2018 Australia*. Australian Government.
Britton, J. C., Lissek, S., Grillon, C., Norcross, M. A., & Pine, D. S. (2011). Development of anxiety: The role of threat appraisal and fear learning. *Depression and Anxiety, 28*, 5–17.
Ferris, J., Mazerolle, L., King, M., Bates, L., Bennett, S., & Devaney, M. (2013). Random breath testing in Queensland and Western Australia: Examination of how the random breath testing rate influences alcohol related traffic crash rates. *Accident Analysis & Prevention, 60*, 181–188.

Freeman, J., Kaye, S.-A., Truelove, V., & Davey, J. (2017). Is there an observational effect? An exploratory study into speed cameras and self-reported offending behavior. *Accident Analysis & Prevention, 108*, 201–208.
Freeman, J., Parkes, A., Lewis, N., Davey, J. D., Armstrong, K. A., & Truelove, V. (2020). Past behaviors and future intentions: An examination of perceptual deterrence and alcohol consumption upon a range of drink driving events. *Accident Analysis & Prevention, 137* 105428.
Freeman, J., Szogi, E., Truelove, V., & Vingilis, E. (2016). The law isn't everything: The impact of legal and non-legal sanctions on motorists' drink driving behaviors. *Journal of Safety Research, 59*, 53–60. https://doi.org/10.1016/j.jsr.2016.10.001.
Freeman, J., & Watson, B. (2006). An application of Stafford and Warr's reconceptualisation of deterrence to a group of recidivist drink drivers. *Accident Analysis & Prevention, 38*, 462–471.
Freeman, J., & Watson, B. (2009). Drink driving deterrents and self-reported offending behaviors among a sample of Queensland motorists. *Journal of Safety Research, 40*, 113–120.
Grosvenor, D., Toomey, T. L., & Wagenaar, A. C. (1999). Deterrence and the adolescent drinking driver. *Journal of Safety Research, 30*, 187–191.
Gual, A., Segura, L., Contel, M., Heather, N., & Colom, J. (2002). Audit-3 and audit-4: effectiveness of two short forms of the alcohol use disorders identification test. *Alcohol and alcoholism (Oxford, Oxfordshire), 37*(6), 591–596. https://doi.org/10.1093/alcalc/37.6.591.
Homel, R. (1988). *Policing and punishing the drinking driver: A study of general and specific deterrence*. Macquarie University.
Lebel, C., & Beaulieu, C. (2011). Longitudinal development of human brain wiring continues from childhood into adulthood. *Journal of Neuroscience, 31*, 10937–10947.
Lenk, K. M., Nelson, T. F., Toomey, T. L., Jones-Webb, R., & Erickson, D. J. (2016). Sobriety checkpoint and open container laws in the United States: Associations with reported drinking-driving. *Traffic Injury Prevention, 17*, 782–787.
Meeter, M., Shohamy, D., & Myers, C. E. (2009). Acquired equivalence changes stimulus representations. *Journal of the Experimental Analysis of Behavior, 91*, 127–141.
Morrison, C. N., Ferris, J., Wiebe, D. J., Peek-Asa, C., & Branas, C. C. (2019). Sobriety checkpoints and alcohol-involved motor vehicle crashes at different temporal scales. *American Journal of Preventive Medicine, 56*, 795–802.
Recio, S. A., Iliescu, A. F., & de Brugada, I. (2019). The amount of exposure determines generalization in animal perceptual learning using short inter-stimulus intervals. *Behavioural Processes, 166* 103900.
Sanjuan, M. D. C., & Nelson, J. B. (2019). Perceptual learning after test-stimulus exposure in humans. *Behavioural Processes, 162*, 1–6.
Szogi, E., Darvell, M., Freeman, J., Truelove, V., Palk, G., Davey, J., & Armstrong, K. (2017). Does getting away with it count? An application of Stafford and Warr's reconceptualised model of deterrence to drink driving. *Accident Analysis & Prevention, 108*, 261–267.
Von Hirsch, A., & Colloquium, C. U. I. O. C. (1999). *Criminal deterrence and sentence severity: An analysis of recent research*. Hart Publishing Oxford.
Watson, A., Freeman, J., Imberger, K., Filtness, A. J., Wilson, H., Healy, D., & Cavallo, A. (2017). The effects of licence disqualification on drink-drivers: Is it the same for everyone? *Accident Analysis & Prevention, 107*, 40–47.
Watson, B. (1994). Enhancing the effectiveness of RBT in Queensland. In *Proceedings of the conference on alcohol-related road crashes: social and legal approaches, 1994* (pp. 31–49). Centre for Crime Policy and Public Safety, Griffith University.
Watson, B., & Freeman, J. (2007). Perceptions and experiences of random breath testing in Queensland and the self-reported deterrent impact on drunk driving. *Traffic Injury Prevention, 8*, 11–19.
Watson, B., Hart, S., Freeman, J., Tay, R., Davey, J., Madden, J., & Goff, A. (2005). *Review of random breath testing (RBT) in Queensland*. Queensland University of Technology.

A/Professor James Freeman works as a Principal Research Fellow at the University of the Sunshine Coast in the USC Road Safety Research Collaboration. He is a registered psychologist with the Australian Health Practitioner Regulation Agency (AHPRA) and is a full member of the Australian Psychological Society (APS) with forensic psychology endorsement. He has over 130 peer review publications, has been awarded a range of competitive national grants and has particular research intersts in perceptions of risk and risk assessment.

Alex Parkes has completed Master of Professional Psychology, Post Graduate Diploma in Psychology and a Bachelor of Psychological Science. Alex currently works as a research officer with the University of the Sunshine Coast's Road Safety Research Collaboration. In this Role, Alex manages the day to day logistics of road safety research including ethics approval, data collection, data analysis and contribution to the publication of research manuscripts. Alex's focus area of research is the impact of substance impairment on severe and fatal road crashes in Queensland Australia.

Verity Truelove is a Postdoctoral Research Fellow at the University of the Sunshine Coast's Road Safety Research Collaboration and has recently completed her PhD at the Centre for Accident Research and Road Safety – Queensland at Queensland University of Technology. Her research focusses on deterring drivers from engaging in risky road rule violations.

Ms Naomi Lewis has completed a Bachelor of Social Science (Psychology) (Honours) degree at USC and was a research assistant for the Road Safety Research Collaboration.

Jeremy Davey (PhD) is a Professor in the School of Social Sciences and Director of the University of the Sunshine Coast, Road Safety Research Collaboration. Jeremy was previously Professor and co-founder of the State's first dedicated university research centre in road safety (CARRS-Q at the Queensland University of Technology). He is a recent awardee of the United States government's flagship foreign exchange scholarship program award of a Senior Fulbright Scholar to undertake comparative road safety research in the US.

Journal of Safety Research 76 (2021) 228–237

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Driver glare exposure with different vehicle frontlighting systems

Patricia Alcón Gil [a], César De Santos-Berbel [b], Maria Castro [c,*]

[a] Universidad Politécnica de Madrid, Spain
[b] Departamento de Estructuras y Física de Edificación, Universidad Politécnica de Madrid, Spain
[c] Departamento de Ingeniería del Transporte, Territorio y Urbanismo, Universidad Politécnica de Madrid, Spain

A R T I C L E I N F O

Article history:
Received 30 May 2020
Received in revised form 18 August 2020
Accepted 22 December 2020
Available online 8 January 2021

Keywords:
Road safety
Nighttime driving
Headlights
3D modeling
Vegetation barrier

A B S T R A C T

Introduction: Highway safety performance at night has received less attention in research than daytime, despite the higher accident rates occurring under night-time conditions. This study presents a procedure to assess the potential hazard for drivers created by headlight glare and its interaction with the geometric design of highways. *Method:* The proposed procedure consists of a line-of-sight analysis performed by a geoprocessing model in geographic information systems to determine whether the rays of light that connect headlights and oncoming drivers are obstructed by either the roadway or its roadsides. Then, the procedure checks whether the non-obstructed rays of light are enclosed by a given headlight beam. Different hypotheses were set concerning the headlight beam features, including the horizontal spread angle and whether the headlights are fixed or swiveling. A highway section was selected to test and validate the procedure proposed. A 3D recreation of the highway and its environment derived from a LiDAR point cloud was used for this purpose. *Results:* The findings disclose how glare is produced on tangents, horizontal curves, transitions between them and sequences of curves. The effect of visual obstructions conveniently placed is also discussed. *Conclusions:* A greater glare incidence is produced as the horizontal headlights spread angle increases. Swiveling headlights increase glare on highways left curves and reduce it on right curves. *Practical Applications:* The procedure and conclusions of this study can contribute to develop more effective glare avoidance technologies as well as identify and assess glare-prone sections. The glare evaluation assists in evaluating glare countermeasures such as deciding whether to place a vegetation barrier and where.

1. Introduction

In nighttime driving, road users must have an adequate view of the roadway and the roadsides without the setback of glare. Vehicle frontlighting systems may dazzle drivers traveling on the opposing direction, thus reducing overall visibility (disability glare), as well as causing distraction and annoyance (discomfort glare) (Hwang & Peli, 2013). Even if a light source ceases to dazzle, it takes some time for the driver's eye to adapt to the variation in light intensity. As a result, the associated contrast reduction may affect drivers' visual performance, potentially affecting highway safety.

Road accidents represent a considerable loss of human life as well as a negative impact on economic activity. Fewer studies have been conducted on nighttime sight distance than on daylight, despite the higher accident rates occurring under nighttime conditions (AASHTO, 2018). Particularly, quantifying the real impact of glare on driving performance is difficult. In fact, accident records can hardly gather such information. However, several authors have described the increased risk of nighttime accidents associated to disability glare caused by vehicle headlights (Babizhayev, 2003; Lachenmayr, Berger, Buser, & Keller, 1998). Notwithstanding the foregoing, vehicle lighting systems have experienced significant advances in recent years and show potential contributions to improve safety performance (Mehler, Reimer, Lavalliere, Dobres, & Coughlin, 2014; Peña-García, Peña, Espín, & Aznar, 2012).

The geometry of the highways also plays an important role in the incidence of headlight glare since certain alignment sequences might contribute to glare. In this sense, a vegetation barrier is presented as a possible treatment to reduce headlight glare. However, its effect has not been studied in 3D. In addition, advanced frontlighting systems should be able to adapt to the existing alignment sequences and traffic to prevent glare occurrence.

The aim of this study is to propose a procedure to observe the incidence of the driver glare caused by vehicle headlights, and evaluate the effect of different headlight features on the driver glare occurring on certain highway alignment sequences and var-

* Corresponding author.
E-mail address: maria.castro@upm.es (M. Castro).

https://doi.org/10.1016/j.jsr.2020.12.018
0022-4375/

ied roadside features such as vegetation. To test the procedure, a real highway was modeled and assessed under a 3D approach.

2. Background

2.1. Drivers glare

Driving performance in nighttime is largely determined by visibility conditions, which are, in turn, influenced by the design of the highway and its environment as well as the headlight features. Drivers must be able to visualize appropriately the roadway ahead sufficiently illuminated while ensuring that the headlights do not disrupt the oncoming traffic.

Glare is a phenomenon that affects two aspects of the drivers' vision (Theeuwes, Alferdinck, & Perel, 2002). On the one hand, disability glare creates reduced contrast sensitivity. On the other hand, discomfort glare produces a sensation of discomfort and fatigue in the driver without loss of vision. To evaluate glare, the 9-point De Boer scale is most widely used in the field of automotive and public lighting (De Boer, 1967). Bullough (2014) analyzed visual performance benefits and quantify potential safety benefits from adaptive high-beam headlamp systems using mesopic and photopic measurements. Van Derlofske, Bullough, Dee, Chen, and Akashi (2004) tested glare placing a light source at 50 m at an angle of 5° to simulate oncoming traffic, measuring the ability of driver to detect a set of targets. Reagan, Frischmann, and Brumbelow (2016) studied the perceived glare produced by vehicles approaching from straight and curved trajectories on either side. The frontlighting systems of the test vehicles included diverse combinations of halogen and high-intensity discharge headlamps, fixed and swiveling, low beams and high beams.

Concerning the separation distance between the driver and the glare vehicle, drivers have been reported to be affected by headlight glare from distances of less than 400 m (Porter, Hankey, & Binder, 2005).

2.2. Headlight features

A tradeoff between appropriate visibility of the driving scene and headlight glare avoidance is necessary. As a result, the photometric design of headlamps is standardized (Gibbons, Medina, Williams, Du, & Rakha, 2012). European headlamp regulations establish stricter limitations for the headlight beam features to prevent glare, in detriment of the visual performance of the own driver. Conversely, American regulations promote higher illuminance performance at the expense of the comfort of oncoming drivers.

Headlamp technology has advanced considerably in recent years and new designs could generate new glare scenarios. In this respect, there is concern that downsizing headlight systems may be contributing to glare. Moreover, headlamp types such as high intensity discharge, halogen and LED, the spectrum and size of the headlamps are factors that affect the severity of glare (Akashi, Hu, & Bullough, 2008; Van Derlofske et al., 2004).

2.3. Nighttime visibility and glare simulation on highways

To simulate the conditions under which headlight glare is produced, a line-of-sight analysis must be undertaken. This analysis concerns the classification of whether two positions in space can be connected by a straight line that founds no obstruction between its ends. A ray of light is a particular case of line of sight, which is assumed to be enclosed within the headlight beam. Glare may therefore be produced if no obstruction is found between the headlamp and the eye of the oncoming driver.

The headlight beam can be modeled in the 3D space as per the volume enclosed by theoretical boundaries given by the positions of the headlights with respect to the driver, the horizontal spread angle, the vertical spread angle, and the range (Hassan, Easa, & Abd El Halim, 1997). This procedure enabled the assessment of highway alignment in nighttime driving. In recent studies, the impact of the headlight beam features on headlight sight distance under such approaches was quantified (De Santos-Berbel, Castro, & Iglesias, 2016). Moreover, variations of the vertical spread angle have been found to potentially produce a significantly different impact on safety (Andrade-Cataño, De Santos-Berbel, & Castro, 2020).

Adaptive curve lighting, and particularly swiveling headlights, involves bending the beam pattern into the curve. The swiveling angle of headlights is the angle rotated by a frontlighting device to bend the lights apart from the tangential axis of the vehicle's trajectory. The value of this angle is typically governed by the turning angle of the vehicle and the speed (Ishiguro & Yamada, 2004). Several experiments carried out in test tracks to assess the effectiveness of swiveling headlights can be found in literature (Bullough et al., 2006, 2016; Hagiwara et al., 2007, 2009; Ishiguro & Yamada, 2004). In general, these lighting systems have been reported to enhance visualization, target detection, and driving performance.

A number of studies characterized the effects of swiveling headlights on horizontal curves through the evaluation of the horizontal projection of the beam area through computer-aided simulations (Gao & Li, 2014; Sivak, Schoettle, Flannagan, & Minoda, 2005). Also, the combined effect of swiveling headlamps and alignment sequences on headlight sight distance has been modeled in 3D (De Santos-Berbel & Castro, 2020).

Situations where headlight glare is produced have been analyzed in diverse simulation systems. Akashi et al. (2008) simulated the photometric distributions of diverse vehicle headlamp types to recreate the glare produced by oncoming traffic on tangents. It was found that higher mounting heights produce greater levels of glare than expected. Moreover, the effect of bright headlights has been reproduced and incorporated into a driving simulator (Haycock, Campos, Koenraad, Potter, & Advani, 2019; Hwang & Peli, 2013). Recent studies have incorporated nighttime driving conditions into a driving simulator to evaluate the driver behavior. Driving simulators have also been utilized for advances in glare-free frontlighting systems (Berssenbrügge, Trächtler, & Schmidt, 2016).

3. Modelling of headlight beam and glare

The research hereby presented was based on previous developments of the authors. The procedure followed is outlined in Fig. 1, which consists of two main components. The first one is the geoprocessing model, which was exploited for the evaluation of lines of sight (Iglesias, Castro, Pascual Gallego, & De Santos-Berbel, 2016). This computational tool operates on geographic information systems (GIS), requiring as inputs a digital terrain model (DTM), a 3D object file, and a dataset containing the points that represent the ends of the lines of sight to be evaluated.

The DTM and the 3D object file were derived from a Light detection and ranging (LiDAR) point cloud collected by a mobile mapping system. LiDAR surveys are a reliable data source to model a highway and its environment for the purpose of sight distance studies (De Santos-Berbel & Castro, 2018; Jung, Olsen, Hurwitz, Kashani, & Buker, 2018; Ma, Zheng, Cheng, & Easa, 2019). The DTM illustrates the shape of the roadway and the terrain that surrounds it. This dataset represents 2.5-D features as each position on the horizontal projection corresponds to a single elevation value, not enabling overhanging features (De Santos-Berbel,

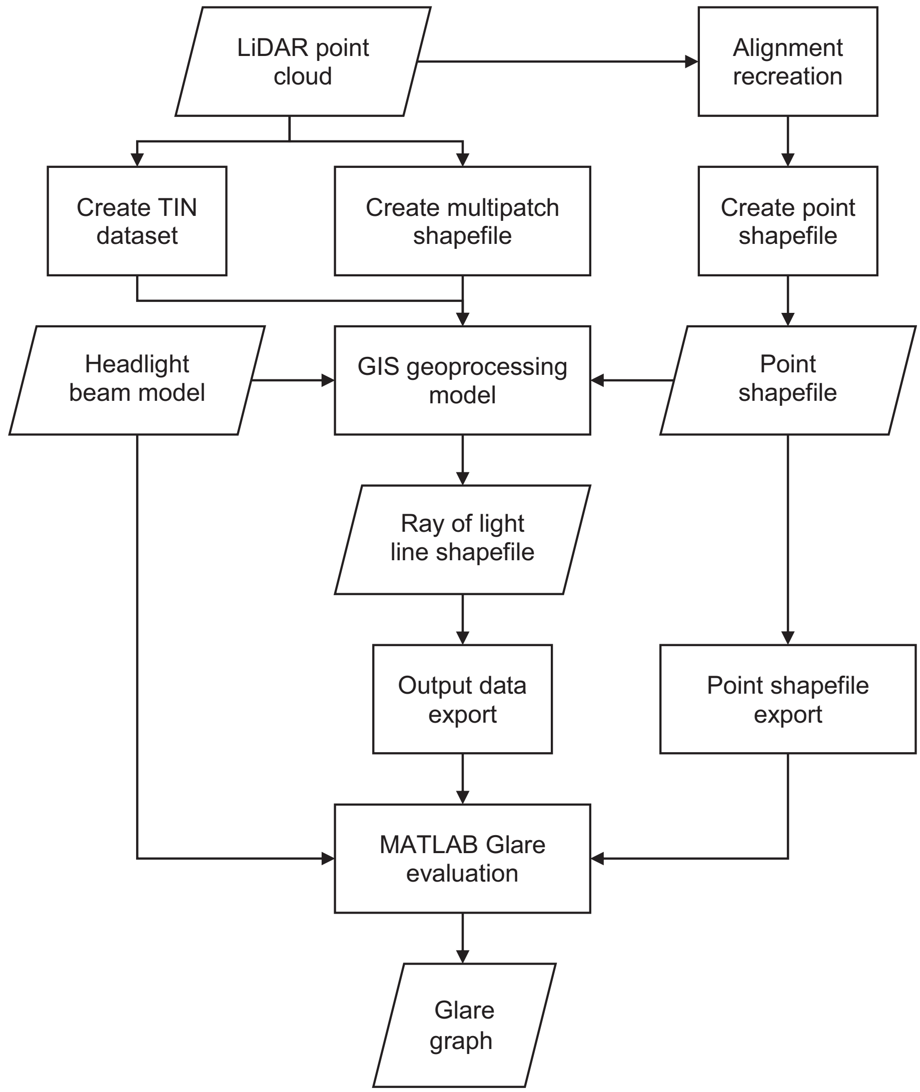


Fig. 1. Flowchart of the procedure for glare evaluation.

Castro, López-Cuervo, & Paréns-González, 2014). Therefore, to connect terrain points of the LiDAR point cloud, a triangular irregular network (TIN) dataset was built up in GIS. The 3D object file was created from the remaining points of the LiDAR point cloud to represent features above the ground with fully-3D features. It connects points closer than a threshold distance and classified within the same class (i.e., vegetation, roadside equipment, etc.) to form polyhedral entities (Arranz Justel, 2013; Iglesias, De Santos-Berbel, Pascual, & Castro, 2019). These features were imported in GIS as a georeferenced multipatch shapefile, which is a GIS object that stores a collection of patches (generally triangular) to represent the boundary of a number of 3D objects (ESRI, 2008). The TIN dataset and the multipatch shapefile constitute the 3D roadway and roadside model itself. Although it is not part of the 3D model itself, a point shapefile representing the ends of the lines of sight to be launched is required. Therefore, this shapefile must contain, on the one hand, points located on the sequence of the headlamps' positions of a vehicle traveling along the highway, where the lines of sight representing rays of light originate. On the other hand, the eye position of the driver of the oncoming traffic, where the lines of sight end, must be included in the point shapefile. This shapefile was derived from the theoretical highway centerline. Both the horizontal and the vertical projections of the alignment were deduced with computer-aided design software. According to the Spanish geometric design standard, the driver is assumed to follow a trajectory parallel to the roadway centerline on the own lane at an offset of 1.5 m (d_1 in Fig. 2) (Ministerio de Fomento, 2016). The position of either headlight in space was set in relation to that of the driver as per the following four geometric parameters (Fig. 2): the headlamp mounting height (h_h), the headlamp headway with respect to the driver (d_2), the offset between headlamps (d_3) and the offset of left headlamp with respect to the driver (d_4). The values selected for these parameters were derived from those of the 11 most sold vehicles in Spain, which are shown in Table 1 (ANFAC, FACONAUTO & GANVAM, 2016). In this sense, it must be noted that the glare evaluation assumed that both vehicles involved are passenger cars.

The stations that represent the vehicle trajectories were obtained by means of a script that calculated points on the centerline spaced 5 meters apart. From these points, stations were obtained at the standardized offset d_1 on both sides of the centerline and, from each station, the counterpart headlamp positions in space. Finally, two point shapefiles were produced: one that corresponds to the outward driving direction, which comprises the headlamp points in such a direction and the driver's eye positions on the return direction; and another one that corresponds to the

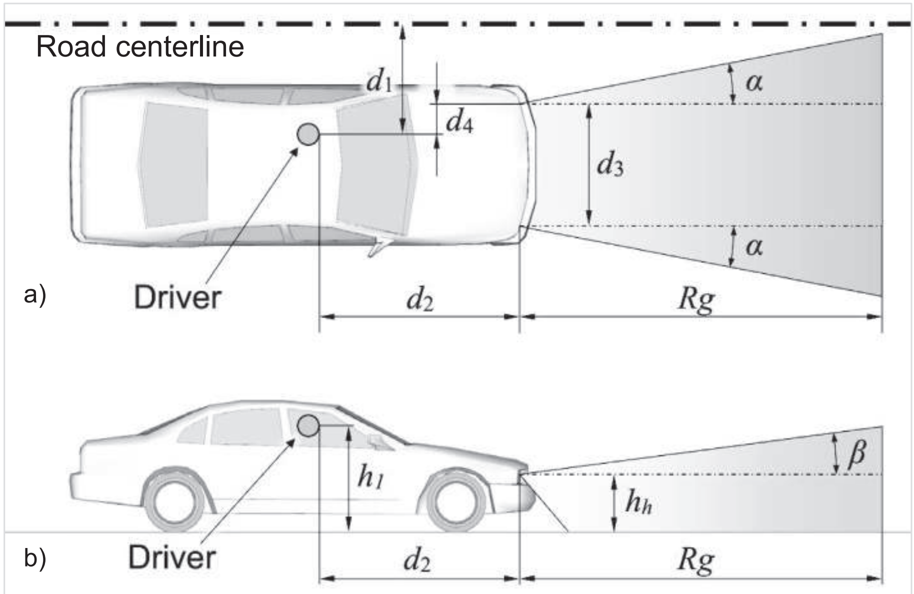


Fig. 2. Layout of headlight beam and driver a) horizontal projection, b) vertical projection.

Table 1
Parameters featuring the position of headlamps.

Parameter	Value
h_h	0.750 m
d_1	1.775 m
d_2	1.345 m
d_3	0.320 m

return driving direction, which contains the headlamp points of the return direction and the driver's eye positions on the outward direction. Fig. 3 displays a section of highway where the point shapefile contains both a sequence of paired positions of the headlights ($h_{i,r}$ and $h_{i,l}$), and a sequence of the driver's eye positions (d_{i+1}, ..., d_{i+6}). The blue lines represent rays of light launched from the headlights at a generic position i towards the positions of the driver's eyes ahead.

The geoprocessing model output a line shapefile connecting the headlamp positions with the driver's eye positions. Each line represents a ray of light and is associated with the line-of-sight evaluation outcome (true or false).

The second component consists of a MATLAB module that is fed with both the output of the geoprocessing model and the geometric parameters of a headlight beam to compute the glare incidence. This module evaluates whether the unobstructed rays of light con-

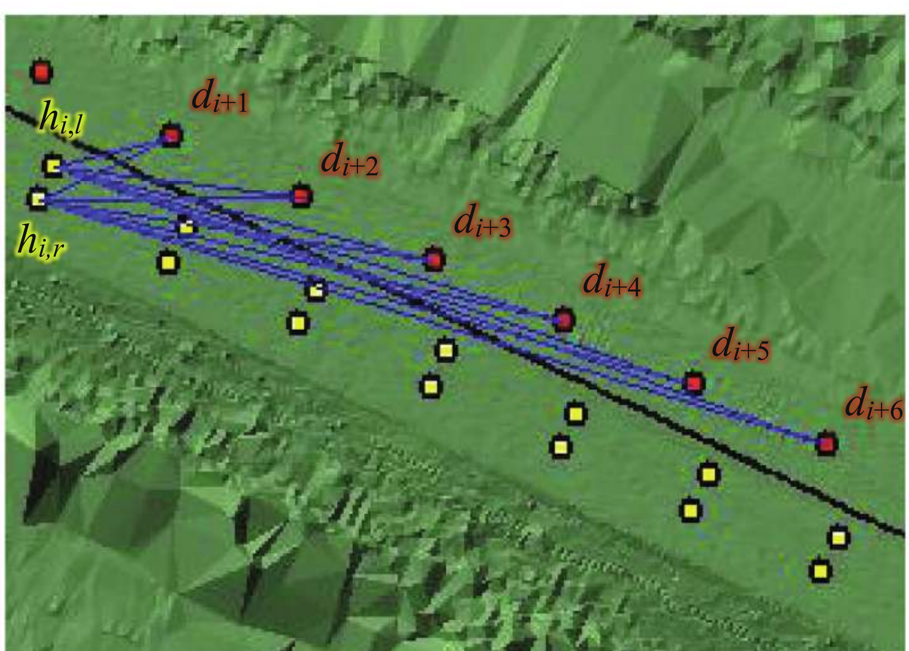


Fig. 3. Schema of lines launched between headlamp positions (yellow spots) and oncoming driver's eyes (red spots).

necting a headlamp position with a target are enclosed within the defined headlight beam. Particularly, the illuminating headlights of a vehicle at a given position are considered to dazzle an oncoming driver at another particular position ahead if, at least, one of the unobstructed rays of light connecting the positions in the 3D space of the headlamps and the driver's eye laid within the headlight beam (Fig. 4). The output of this module can be displayed in a glare graph. The glare graph is a plot that features headlight glare incidence along a highway section. The abscissa axis indicates the stations where the vehicle considered to be illuminating is, and the ordinate axis specifies the distance along the centerline to a position ahead that may be affected by glare.

For the evaluation of glare, the headlight beam layout adopted is based on assumptions made in previous studies (De Santos-Berbel & Castro, 2020; De Santos-Berbel et al., 2016; Hassan et al., 1997). It is assumed to be bounded by the horizontal spread angle (α), the upward vertical spread angle (β) and the headlight glare range (Rg). A graphical description of these parameters is found in Fig. 2. With regard to the horizontal spread angle, Hassan et al. (1997) explored horizontal spread angle values ranging from 0° to 10° outward the spotlights in headlight sight distance evaluations. Later studies have utilized values ranging from 0° to 6° (De Santos-Berbel & Castro, 2020; De Santos-Berbel et al., 2016). The vertical spread angle is measured upward the beam axle, which is in turn parallel to the longitudinal grade of the roadway. A value of 1° is commonly found in standards, guides and research studies (AASHTO, 2018; De Santos-Berbel & Castro, 2020; Hassan et al., 1997; Ministerio de Fomento, 2016). The roadway grade was accurately extracted from the LIDAR surveying data. To analyze the sensitivity of headlight glare to the variation of the horizontal spread angle, values between 1° and 6° were assigned to this parameter. Finally, based on the literature, the selected headlight glare range was 400 m. The selected values for the headlight beam parameters are presented in Table 2.

Two types of frontlighting systems were regarded: fixed frontlighting systems (FFS) and swiveling frontlighting systems (SFS). The consideration of different types of frontlighting systems involved the assumption of a certain swiveling angle (φ in Fig. 4). The value of this angle remained constant and equal to zero along the highway section in the case of FFS, whereas it varied in the case of SFS. Based on previous studies developed by the authors, the swiveling angle of SFS is assumed to be controlled by the steering wheel rotation for the purpose of keeping the beams on the roadway where it bends (De Santos-Berbel & Castro, 2020). In turn, the steering wheel turning angle is assumed to be directly proportional to the vehicle's yaw variation per traveled length. The vehicle yaw is the angle rotated by the vehicle around its vertical axis changing the direction of heading. If the yaw is denoted by the letter ψ, the vehicle yaw variation per traveled length, or simply the yaw variation θ, is given by:

$$\theta = \frac{d\psi}{ds} \quad (1)$$

Table 2
Parameters featuring the headlight beam.

Parameter	Value					
α	1°	2°	3°	4°	5°	6°
C_{right}	7.1	8.2	9.6	10.7	11.3	12.5
E_{right}	0.79	0.61	0.58	0.54	0.48	0.46
C_{left}	11.75	12.74	14.13	15.22	15.82	17.02
E_{left}	0.649	0.558	0.544	0.52	0.479	0.464
β	1°					
Rg	400 m					

Also according to the abovementioned study developed by the authors, a mathematical function for the headlamp swiveling angle φ was proposed as follows (De Santos-Berbel & Castro, 2020):

$$\varphi = f(\theta) = C \cdot \theta^E \quad (2)$$

where C and E are calibrated parameters, different for either swiveling side. These calibrated values were used in this study and are displayed in Table 2.

4. Case study

A section of a two-lane rural highway located in the region of Madrid (Spain) was selected to assess the incidence of headlight glare using the procedure devised. Its cross section consists of one 3-m wide lane per driving direction and 0.5-m wide shoulders. A design speed of 80 km/h was assumed as per the geometric features. The main horizontal alignment features are presented in Table 3. An initial reverse curve in which both curve radii are 350 m is followed by short tangent, after which a second reverse curve is found, where both radii are 150 m. This highway section was selected because it comprises a sequence of multiple reverse curves with a combination of clear and non-clear inner roadsides, as well as tangents. These features make it suitable to test the procedure hereby presented for evaluating the incidence of headlight glare and how the different factors considered affect glare caused by vehicle headlights.

Concerning the highway and its environment, two scenarios were contemplated to assess the impact of certain roadside features on headlight glare incidence. Firstly, the actual highway environment was modeled in full accordance with the LiDAR point cloud. A 3D perspective of the original recreated scene, along with the driver's eye positions on either path, is illustrated in Fig. 5a. Secondly, the original multipatch shapefile was added a feature consisting of a visual barrier by the inner roadside at a double reverse curve, were headlight glare was expected to be produced since it was initially clear. The visual barrier consisted of a cluster of trees conceived in SketchUp that can be imported into a multipatch shapefile (De Santos-Berbel & Castro, 2018). The location, size, and layout of the visual barrier were set aided by 3D Analyst tools. Particularly, a set of lines of sight was launched between the beginning and the end of the double reverse curve to obtain the

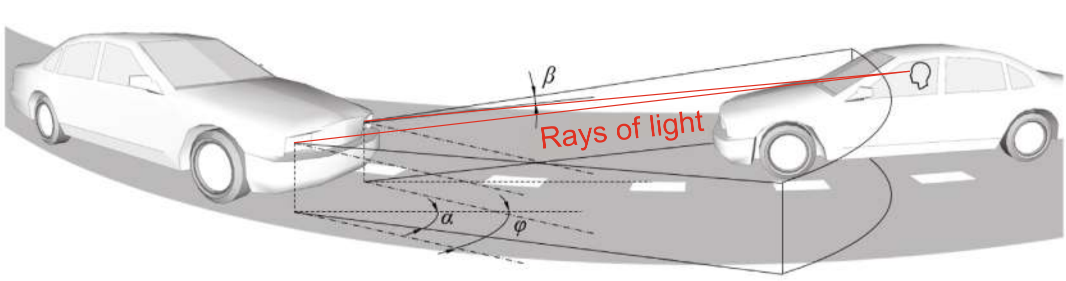


Fig. 4. Layout of headlight beam in 3D space and oncoming driver potentially affected by glare.

Table 3
Characterization of the horizontal alignment of the selected section.

Alignment	Length (m)	Radius (m)	Deflection angle (gon)	Parameter of spiral 1	Parameter of spiral 2
Tangent	556.242	(Inf)	0	–	–
Horizontal curve	225.509	350	27.879	160	158
Horizontal curve	182.558	−350	−17.020	158	193.2
Tangent	17.969	(Inf)	0	–	–
Horizontal curve	131.019	150	33.805	90	85.5
Horizontal curve	136.756	−150	−33.123	85.5	101.5
Tangent	494,465	(Inf)	0	–	–

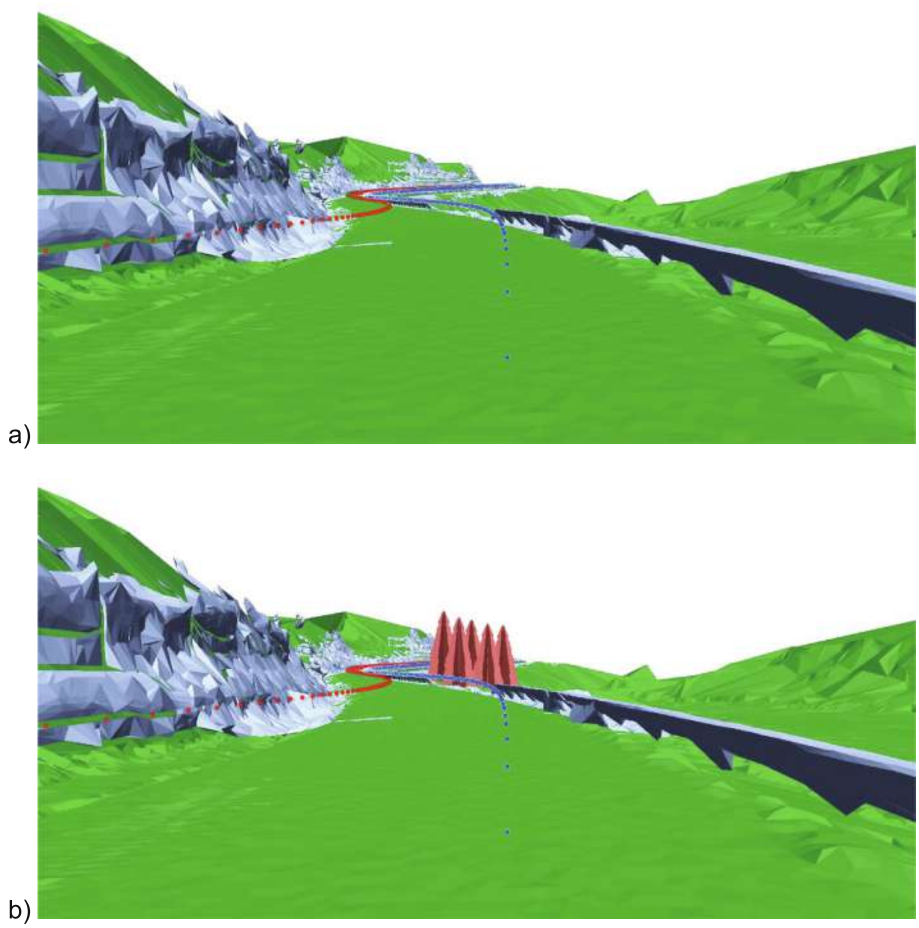


Fig. 5. Scene of the modeled highway a) actual situation and b) with added visual obstruction (cluster of trees).

terrain profile and the necessary heights of the visual barrier. A 3D perspective of the new scene and the driver's eye positions are displayed in Fig. 5b. In the figure, it is observed that the highway section has a composite profile.

5. Results and discussion

In this study, three factors were studied in the 3D simulation of the headlight glare exposure. First, the effect of the horizontal spread angle of light on glare was assessed. Second, the glare produced by two types of frontlighting systems was examined. Third, two scenarios with varied roadside features were analyzed. It must also be noted that glare was evaluated on both driving directions. The results obtained for each of the groups of hypotheses are described and discussed below.

5.1. Effect of the horizontal spread angle on glare

As previously described, headlight glare was assessed on the selected highway section contemplating different values of the horizontal spread angle. To isolate the effect of such a variable, the other factors are kept constant. Fig. 6 displays the headlight glare incidence for FFS on the outward direction in relation to the horizontal spread angle (from 1° to 6°) and the horizontal curvature of the selected section without the additional visual barrier. It is first noticed that glare had more incidence if a wider horizontal spread angle is considered. If the results are observed in relation to the curvature graph, the glare incidence occurred at constant distances on tangents, where the curvature is null. When approaching or while driving through right-hand bends (stations 556 to 782 and 982 to 1,113), oncoming traffic is dazzled at shorter distances than on tangents, especially as the horizontal spread

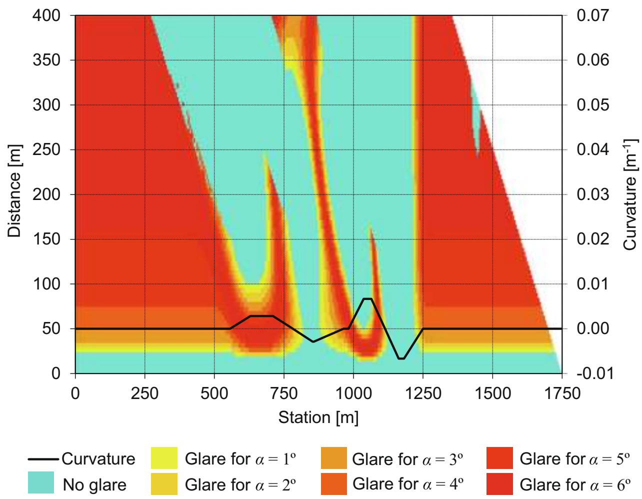


Fig. 6. Combined glare graph for FFS with the horizontal spread angle.

angle increases, but the section affected by headlight glare is far shorter. Conversely, the headlight glare occurrence is virtually non-existing on left-hand bends (stations 782 to 964 and 1,113 to 1,250). It can also be observed that on left-hand bends, the existing cut slope produced a visual barrier that breaks the glare area on the graph around station 750.

These results can help estimate the duration of glare when the vehicles involved are travelling at a given speed. Thus, assuming that both vehicles travel at the design speed (80 km/h) alongside the tangent, when glare occurs, it would last from 5.6 s in the case of $\alpha = 1°$ to 8.4 s in the case of $\alpha = 6°$. Although the duration of the headlight glare would be much shorter in curves, the situation would be significantly more hazardous if the dazzled driver experiences disability glare as they could lose all reference to his/her position on the roadway on a curved alignment.

5.2. Effect of the frontlighting system on glare

As mentioned earlier in this document, the glare produced by two frontlighting systems was considered: FFS and SFS. To analyze the effect of this factor, the headlight glare incidence considering constant horizontal spread angle of 4° on the outward direction without the additional visual barrier is assessed. Fig. 7 illustrates the corresponding glare graph along with the horizontal curvature. On right-hand curves, FFS produce glare for shorter distances between the illuminating vehicle and the oncoming traffic. Conversely, on left-hand curves, SFS increase the incidence of glare, which is practically non-existing for FFS. This effect is produced because the light beam bends toward the inside of the curve to light up the roadway instead of lighting up the roadside that lies straight ahead. In left-hand bends, the light from the swiveling headlamp runs through the opposite lane in order to illuminate the greatest possible distance, dazzling at shorter distances. On the contrary, in right-hand bends, the light from the swiveling headlamp passes over the inner roadside, avoiding glare on short distances. It can also be noted that the minimum distance at which glare is produced on curves (20–30 m) is significantly shorter than the counterpart distance on tangents (40 m).

Additional noteworthy findings concern the relationship between the radius of the horizontal curves and the ranges of glare distances. On the one hand, the minimum distance at which headlight glare is produced by the calibrated SFS was independent from the curve radius on left-hand bends, they resulting as short as 15 m. However, the maximum distance at which the oncoming driver was dazzled was significantly longer on the left-hand bend around station 680 than on the left-hand bend around station 1,180. This difference was likely to be produced by the small deflection angle turned on the former curve. On the other hand, the minimum distances at which glare was produced when assessing the FFS decreased as the curve radius increased.

Fig. 8 shows the percentage of stations as a function of the distances ahead where headlight glare was found to be produced, in total for both driving directions. Three pairs of series that corre-

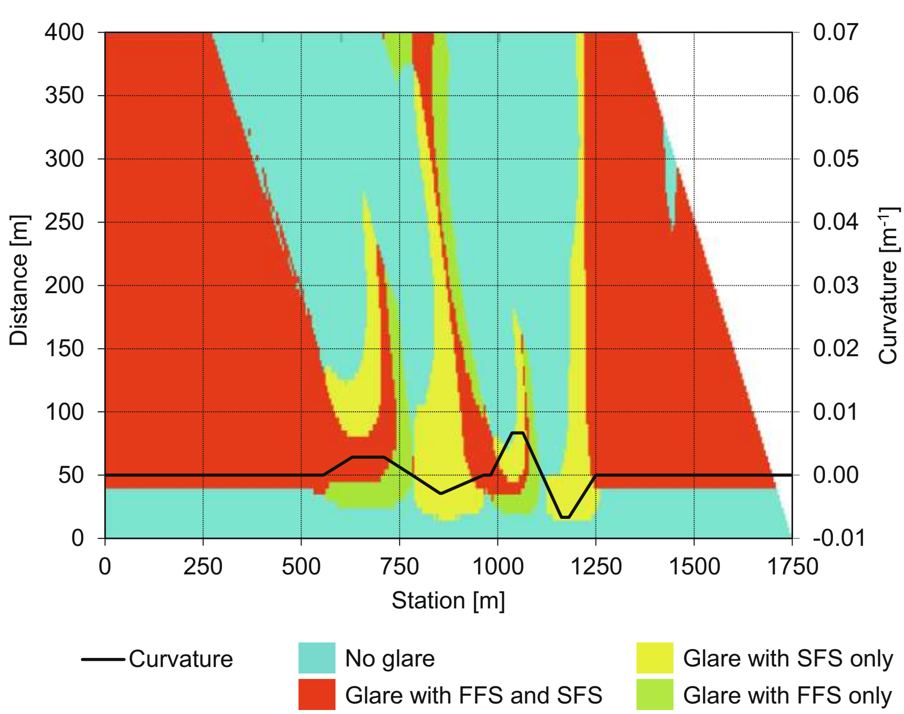


Fig. 7. Combined glare graph for $\alpha = 4°$ with FFS and SFS.

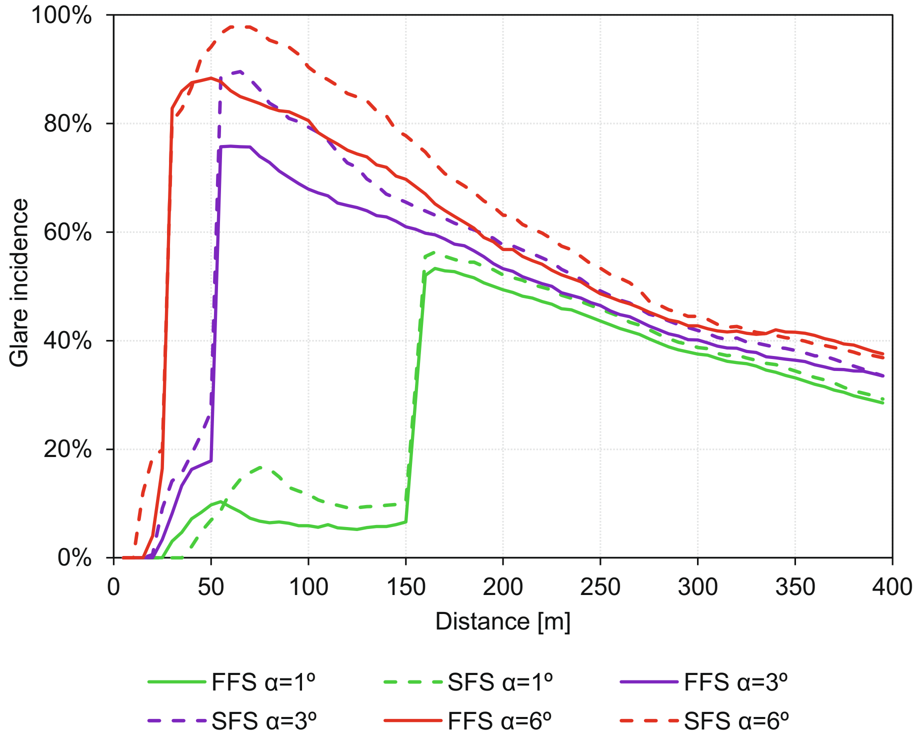


Fig. 8. Overall percentage of glare incidence as a function of the distance.

spond to three horizontal spread angle values (1°, 3° and 6°) in combination with the two frontlighting systems considered in this study are represented. It is first noticed that all the series present a sudden increase of the headlight glare incidence percentage at different distances. As the horizontal spread angle increases, the distance at which the peak is found decreases. These values correspond to the distance from which glare is produced on tangents as deduced from Fig. 6. It can also be observed that the SFS generally yielded slightly higher percentages of glare incidence at all ranges of distances. These results highlight the increased tendency of SFS to produce glare.

5.3. Effect of the additional visual barrier

As indicated earlier, the effect of introducing a visual barrier by the inner roadside of a double reverse curve on headlight glare was assessed. The glare outcome produced in the two scenarios is illustrated in Fig. 9 for its comparison. To analyze this effect, FFS with

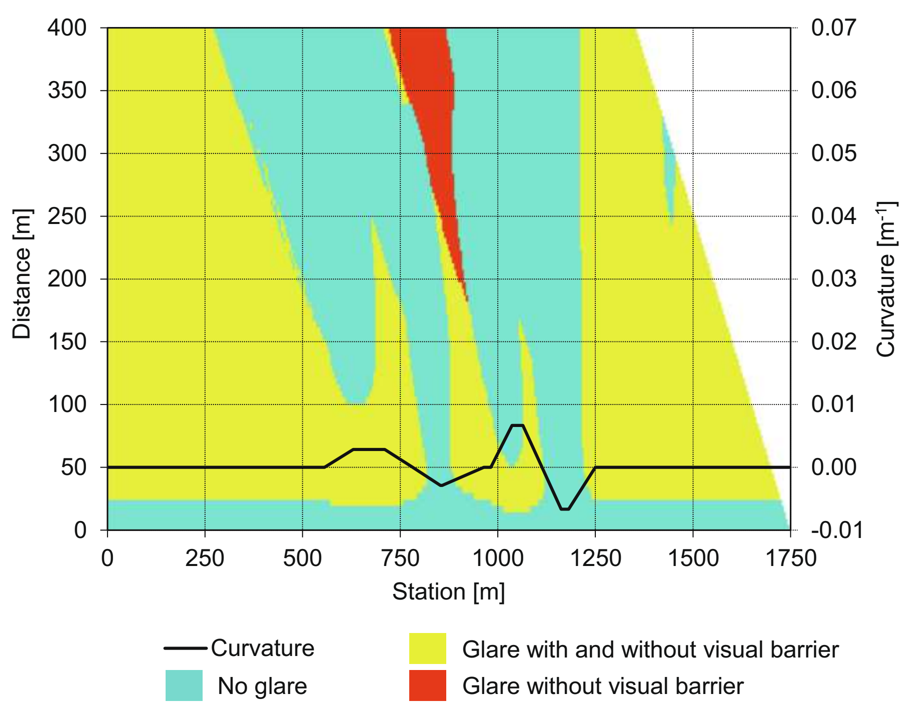


Fig. 9. Combined glare graph for α = 6° and FFS with the effect of visual barrier.

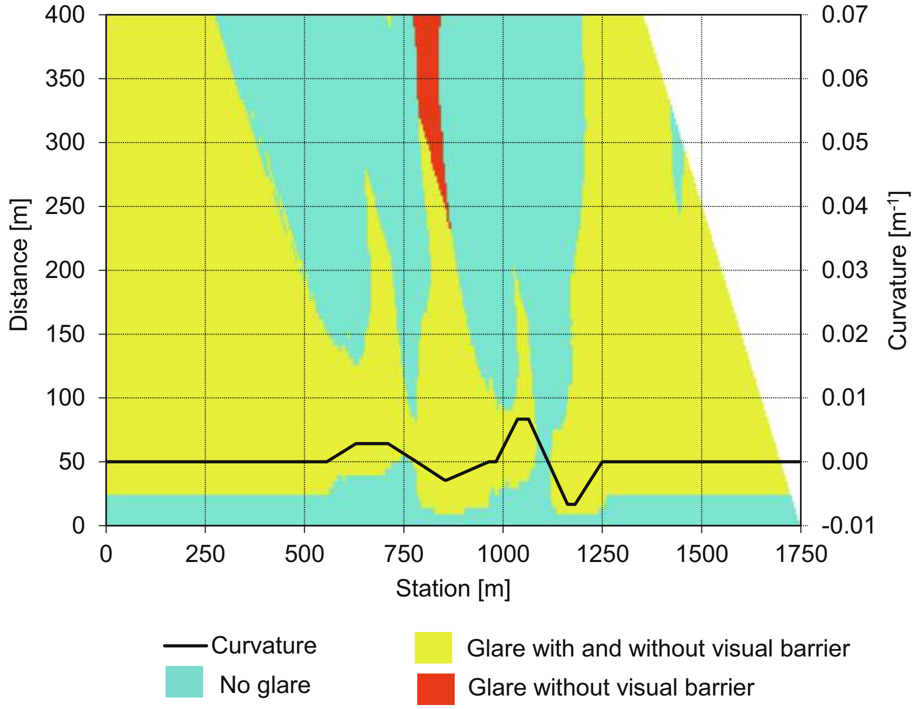


Fig. 10. Combined glare graph for α = 6° and SFS with the effect of visual barrier.

the horizontal spread angle set at 6° were considered on the outward direction. In this direction, the visual barrier is located between the two left-hand curves. Its effect, highlighted in red in the figure, is therefore produced when the dazzling vehicle travels along the first left-hand curve, and affects targets located on the second left-hand curve. Moreover, given the existing sequence of alignments, the sections where glare is avoided are at a distance of more than 200 m.

Finally, a similar comparison was performed considering SFS with the horizontal spread angle set at 6° on the outward direction (Fig. 10). It can be noticed that the situations where headlight glare was produced, in general, increased in comparison to the FFS as the yellow area is greater than in Fig. 9. However, the impact of the visual barrier on headlight glare reduction is less significant in the case of the SFS. This occurred because the SFS produce less glare between vehicles on the two left-hand bends than FFS do. Instead, glare is produced between vehicles located on consecutive curves if SFS are regarded.

6. Conclusions

This study proposes a procedure to address the potential highway safety effects of headlight glare and their interaction with the geometric design of a highway section. A 3D model of the highway and its environment was derived from a LiDAR point cloud. A line-of-sight analysis performed by a geoprocessing model in GIS determined whether the rays of light emitted by headlights hit the eye of the oncoming driver or are intercepted by either the roadway or the roadside obstructions. Then, a MATLAB module was developed to check whether the non-obstructed rays of light are enclosed by a given headlight beam. Different hypotheses were set concerning the headlight beam features, including the horizontal spread angle and two types of frontlighting systems, namely FFS and SFS. A highway section was selected to test and validate the procedure hereby proposed.

The results showed that a greater glare incidence is produced on oncoming drivers as the horizontal spread angle of the headlight beam increases. It was also found that SFS increased glare occurrence on left-hand curves, while glare incidence was reduced on right-hand bends when compared to the expected glare of FFS. However, a more efficient orientation of the light beam should not increase the incidence of glare. The results also indicated that SFS are overall more prone to produce glare than the FFS.

A vegetation virtual visual barrier was conveniently placed by the roadside of an existing double reverse curve on the highway model, and its effect on headlight glare avoidance was studied. Given the highway geometric layout of the case study presented, its effectiveness on glare avoidance was noticed for distances beyond 200 m. A headlight glare evaluation assists in deciding whether to place a vegetation barrier and where and can help determine its potential benefits. In addition, it helped validate the procedure hereby proposed as the effects found in the case study were in line with the expectations. The procedure hereby presented can, on the one hand, contribute to develop more effective glare avoidance technologies. On the other hand, it facilitates the identification of glare-prone sections and helps estimate how long glare would last as a measure to assess potential alignment shortcomings. It can also assist in discerning whether headlight glare might have contributed to the occurrence of a particular accident. Moreover, it is capable of evaluating the effectiveness of countermeasures such as a visual barrier.

Given the versatility of the procedure, the authors propose the analysis of glare incidence considering other vehicle types such as heavy vehicles as a future line of research for a more comprehensive evaluation. In addition, the evaluation of glare on a driving simulator recreating the highway segment studied and the light conditions created by the swiveling headlights is to be pursued.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the Ministerio de Economía y Competitividad of Spain and the European Regional Development Fund (FEDER) [grant number TRA2015-63579-R (MINECO/FEDER)].

References

Akashi, Y., Hu, F., & Bullough, J.D. (2008). Sensitivity analysis of headlamp parameters affecting visibility and glare. (No. DOT HS 811 055). Springfield, VA.

American Association of State Highway and Transportation Officials (AASHTO) (2018). *A policy on geometric design of highways and streets* (7th ed.). Washington, D.C: AASHTO.

Andrade-Cataño, F., De Santos-Berbel, C., & Castro, M. (2020). Reliability-based safety evaluation of headlight sight distance applied to road sag curve standards. *IEEE Access, 8*, 43606–43617. https://doi.org/10.1109/ACCESS.2020.2977258.

Arranz Justel, J. J. (2013). *Diseño, optimización y análisis de sistemas basados en técnicas láser, para el modelado geométrico, registro y documentación, aplicados a entidades de interés patrimonial* Doctoral thesis. Universidad Politécnica de Madrid.

Asociación Española de Fabricantes de Automóviles y Camiones (ANFAC), Concesionarios de Automoción (FACONAUTO) Asociación Nacional de Vendedores de Vehículos a Motor Reparación y Recambios (GANVAM), 2016. Matriculaciones de vehículos de turismo 2015. Madrid, Spain.

Babizhayev, M. A. (2003). Glare disability and driving safety. *Ophthalmic Research, 35*, 19–25. https://doi.org/10.1159/000068199.

Berssenbrügge, J., Trächtler, A., & Schmidt, C. (2016). Visualization of headlight illumination for the virtual prototyping of light-based driver–assistance systems. *Journal of Computing and Information Science in Engineering, 16*. https://doi.org/10.1115/1.4033228 030901.

Bullough, J.D. (2014). Adaptive high beam systems: Visual performance and safety effects. SAE Technical Papers Series 2014-01-04. https://doi.org/10.4271/2014-01-0431.

Bullough, J.D., Skinner, N.P., & Plummer, T.T. (2016). Assessment of adaptive driving beam photometric performance. SAE Technical Papers Series 2016-01-14. https://doi.org/10.4271/2016-01-1408.

Bullough, J.D., Van Derlofske, J., & Akashi, Y. (2006). Strategies for optimizing headlamp illumination and visibility along curves in driver vision and lighting; rear vision and indirect vision. SAE Technical Papers Series 2006-01-04.

De Boer, J. B. (1967). Visual perception in road traffic and the field of vision of the motorist. In J. B. De Boer (Ed.), *Public lighting* (pp. 11–96). Eindhoven, Netherlands: Philips Technical Library.

De Santos-Berbel, C., & Castro, M. (2020). Effect of vehicle swiveling headlamps and highway geometric design on nighttime sight distance. *Mathematics and Computers in Simulation, 170*, 32–50. https://doi.org/10.1016/j.matcom.2019.08.012.

De Santos-Berbel, C., & Castro, M. (2018). Three-dimensional virtual highway model for sight-distance evaluation of highway underpasses. *Journal of Surveying Engineering, 144*, 05018003. https://doi.org/10.1061/(ASCE)SU.1943-5428.0000258.

De Santos-Berbel, C., Castro, M., & Iglesias, L. (2016). Influence of headlamp lighting parameters on nighttime sight distance. In *Proceedings of the 3rd international conference on traffic and transport engineering (ICTTE)* (pp. 88–94).

De Santos-Berbel, C., Castro, M., López-Cuervo, S., & Paréns-González, M. (2014). Sight distance studies on roads: Influence of digital elevation models and roadside elements. *Procedia - Social and Behavioral Sciences, 160*, 449–458. https://doi.org/10.1016/j.sbspro.2014.12.157.

Environmental Science Research Institute (ESRI). (2008). The multipatch geometry type. Redlands, California.

Gao, Z., & Li, Y. (2014). A study of bending mode algorithm of adaptive front-lighting system based on driver preview behavior. *Mathematical Problems in Engineering, 2014*, 1–11. https://doi.org/10.1155/2014/939218.

Gibbons, R. B., Medina, A., Williams, B., Du, J., & Rakha, H. (2012). *Sag vertical curve design criteria for headlight sight distance*. Blacksburg, VA: NCHRP Web Document 198.

Hagiwara, T., Hamaoka, H., Tabata, Y., & Morishita, M. (2009). Effect of headlight swivel angle on driver behavior to avoid conflict with pedestrians approaching from the right. *Transportation Research Record, 2138*, 102–111. https://doi.org/10.3141/2138-14.

Hagiwara, T., Morishita, M., Horii, Y., Miki, K., & Ohshima, I. (2007). Preferences for headlight swivel angles at curves on rural two-lane highways. *Transportation Research Record, 2030*, 47–53. https://doi.org/10.3141/2030-07.

Hassan, Y., Easa, S. M., & Abd El Halim, A. O. (1997). Modeling headlight sight distance on three-dimensional highway alignments. *Transportation Research Record, 1579*, 79–88. https://doi.org/10.3141/1579-10.

Haycock, B. C., Campos, J. L., Koenraad, N., Potter, M., & Advani, S. K. (2019). Creating headlight glare in a driving simulator. *Transportation Research Part F: Traffic Psychology and Behaviour, 61*, 93–106. https://doi.org/10.1016/j.trf.2017.10.006.

Hwang, A. D., & Peli, E. (2013). Development of a headlight glare simulator for a driving simulator. *Transportation Research Part C: Emerging Technologies, 32*, 129–143. https://doi.org/10.1016/j.trc.2012.09.003.

Iglesias, L., Castro, M., Pascual Gallego, V., & De Santos-Berbel, C. (2016). Estimation of sight distance on highways with overhanging elements. In *Proceedigns of the 3rd international conference on traffic and transport engineering (ICTTE)* (pp. 75–82).

Iglesias, L., De Santos-Berbel, C., Pascual, V., & Castro, M. (2019). Using small unmanned aerial vehicle in 3D modeling of highways with tree-covered roadsides to estimate sight distance. *Remote Sensing, 11*, 1–13. https://doi.org/10.3390/rs11222625.

Ishiguro, K., & Yamada, Y. (2004). Control technology for bending mode AFS. SAE Technical Papers Series 2004-01-04.

Jung, J., Olsen, M. J., Hurwitz, D. S., Kashani, A. G., & Buker, K. (2018). 3D virtual intersection sight distance analysis using lidar data. *Transportation Research Part C: Emerging Technologies, 86*, 563–579. https://doi.org/10.1016/j.trc.2017.12.004.

Lachenmayr, B., Berger, J., Buser, A., & Keller, O. (1998). Reduziertes Sehvermögen führt zu erhöhtem Unfallrisiko im Straûenverkehr. *Der Ophthalmology, 95*, 44–50.

Ma, Y., Zheng, Y., Cheng, J., & Easa, S. (2019). Real-time visualization method for estimating 3D highway sight distance using LiDAR data. *Journal of Transportation Engineering, Part A: Systems, 145*, 1–14. https://doi.org/10.1061/JTEPBS.0000228.

Mehler, B., Reimer, B., Lavalliere, M., Dobres, J., & Coughlin, J. F. (2014). *Evaluating technologies relevant to the enhancement of driver safety*. AAA Foundation for Traffic Safety.

Ministerio de Fomento (2016). *Trazado. Instrucción de Carreteras. Norma 3.1-IC.* Boletín Of. del Estado. https://doi.org/10.1017/CBO9781107415324.004.

Peña-García, A., Peña, P., Espín, A., & Aznar, F. (2012). Impact of Adaptive Front-lighting Systems (AFS) on road safety: Evidences and open points. *Safety Science, 50*, 945–949. https://doi.org/10.1016/j.ssci.2011.10.013.

Porter, R.J., Hankey, J.M., Binder, S.C., & Dingus, T.A.0 (2005). Enhanced night visibility series, Volume VII: Phase II—Study 5: Evaluation of discomfort glare during nighttime driving in clear weather (No. FHWA-HRT-04-138). McLean, VA. https://doi.org/10.1007/s13398-014-0173-7.2

Reagan, I. J., Frischmann, T., & Brumbelow, M. L. (2016). Test track evaluation of headlight glare associated with adaptive curve HID, fixed HID, and fixed halogen low beam headlights. *Ergonomics, 59*, 1586–1595. https://doi.org/10.1080/00140139.2016.1154986.

Sivak, M., Schoettle, B., Flannagan, M. J., & Minoda, T. (2005). Optimal strategies for adaptive curve lighting. *Journal of Safety Research, 36*, 281–288. https://doi.org/10.1016/j.jsr.2005.05.003.

Theeuwes, J., Alferdinck, J. W. A. M., & Perel, M. (2002). Relation between glare and driving performance. *Human Factors, 44*, 95–107. https://doi.org/10.1518/0018720024494775.

Van Derlofske, J., Bullough, J.D., Dee, P., Chen, J., & Akashi, Y. (2004). Headlamp parameters and glare. SAE Technical Papers 2004-01-12. https://doi.org/10.4271/2004-01-1280.

Patricia Alcón-Gil obtained her B.S. degree in Civil and Environmental Engineering and her M.Sc. degree (2019) in Civil Engineering from the Universidad Politécnica de Madrid (Technical University of Madrid). Her research is mainly on highway design and traffic safety.

César de Santos-Berbel is currently Assistant Professor in the Department of Structures and Building Physics. He received his PhD in Civil Engineering Systems from the Universidad Politécnica de Madrid (Technical University of Madrid). Since 2012 he has been engaged in research and development of new techniques to evaluate the geometric design of highways. He is author and co-author of various papers in international conferences and journals.

Maria Castro is an Associate Professor at the Universidad Politécnica de Madrid (Technical University of Madrid). She holds a PhD degree in engineering from the Universidad Politécnica de Madrid and a Civil Engineering Master degree from the same university. She held several civil engineering positions in the private sector for eight years before joining academia. Her research interests include transportation, GPS and GIS applications and engineering education. She is author or co-author of numerous papers in international conferences and journals.

JOURNAL OF SAFETY RESEARCH

Journal of Safety Research 76 (2021) 276–288

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Effect of pavement condition and geometrics at signalised intersections on casualty crashes

Nasreen Hussein [a,*], Rayya Hassan [b,1], Michael T. Fahey [c,2]

[a] Department of Civil and Construction Engineering, Swinburne University of Technology, Victoria, 3122, Australia and University of Duhok, College of Engineering, Civil Department, Kurdistan Region Iraq
[b] Department of Health Science and Biostatistics, Swinburne University of Technology, Victoria 3122, Australia
[c] Department of Health Science and Biostatistics, Swinburne University of Technology, Victoria, 3122, Australia

ARTICLE INFO

Article history:
Received 26 March 2020
Received in revised form 14 September 2020
Accepted 21 December 2020
Available online 28 January 2021

Keywords:
Signalised intersections
Casualty crashes
Generalised Estimating Equation
Pavement condition
Geometric characteristics

ABSTRACT

Introduction: This study investigated the effects of pavement surface condition and other control factors on casualty crashes at signalized intersections. It involved conducting a before and after study for road surface condition and situational factors. It also included assessing the effects of geometric characteristics on safety performance of signalized intersections post resurfacing to control for the effect of pavement surface condition. Pavement surface condition included roughness, rutting, and skid resistance. The control factors included traffic volume, light and surface moisture condition, and speed limit. The geometric characteristics included approach width, number of lanes, intersection depth, presence of median, presence of shared lane, and presence of bus stop. *Method:* To account for the repeated observations of the effect of light and surface moisture conditions in four occasions (day-dry, day-wet, night-dry and night-wet) Generalized Estimating Equation (GEE) with Negative Binomial (NB) and log link function was applied. For each signalized intersection in the sample, condition data are collected for the year before and after the year of surface treatment. Crash data, however, are collected for a minimum of three and maximum of five years before and after treatment years. *Results:* The results show that before treatment, light condition, road surface moisture condition, and skid resistance interaction with traffic volume are the significant contributors to crash occurrence. For after treatment; light condition, road surface moisture condition, their interaction product, and roughness interaction with light condition, surface moisture condition, and traffic volume are the significant contributors. The geometric variables that were found to have significant effects on crash frequency post resurfacing were approach width interactions with presence of shared lane, bus stop, or median. *Conclusions:* The findings confirm that resurfacing is significant in reducing crash frequency and severity levels. *Practical Applications:* The study findings would help for better understanding of how geometric characteristics can be improved to reduce crash occurrence.

1. Introduction

Crash frequencies at signalised intersections could be a reflection of poor pavement surface condition. It is believed that adequate level of pavement surface condition would improve traffic safety at signalised intersections and roadway segments. Therefore, a good understanding of the factors contributing to crash occurrence at signalised intersections is important for improving their safety performance. Pavement condition parameters including skid resistance, roughness and rutting are the most common indicators of road safety problems. The findings from past published studies related to the effects of these condition parameters on safety performance involved using different methodologies and cover different locations (intersections and road links) and operating environments (rural and urban).

Generally, these studies show that one of the most important characteristics of road surface which has the best-established relationship with crash occurrence is skid resistance. A number of studies have shown that an improvement in skid resistance can produce considerable safety benefits for both intersections and road segments (Candappa, Scully, Newstead, & Corben, 2007; Giles & Sabey, 1959; Kinnear, Lainson, & Penn, 1984; Noyce, Bahia, Yambo, & Kim, 2005; Oliver, 1999; Saplioglu, Eriskin,

* Corresponding author at: University of Duhok, College of Engineering, Civil Department, Kurdistan Region_Iraq.

E-mail addresses: nasreen.hussein@uod.ac (N. Hussein), rahassan@swin.edu.au (R. Hassan), mfahey@swin.edu.au (M.T. Fahey).

1 Ph: +61 3 9214 8284, Fax +61 3 9214 8264.

2 Ph: +61 3 92148113.

https://doi.org/10.1016/j.jsr.2020.12.021
0022-4375/

Yuzer, & Aktas, 2012). Also, high skid resistance treatments have also shown to achieve a significant reduction in crash occurrence and the severity of some crashes. In accordance with observations from analysis of data for three years before and 18 months after treatment, Kinnear et al. (1984) reported that there was an efficient reduction in crashes by 25% with a 99% confidence level for all surfaces after treatment.

Another study carried out by Simpson and Eng (2005) in Melbourne and Geelong, Victoria, to investigate the performance of roads with high skid resistance and its effect on crash rate trends. Analysis of the data for before and after study indicated a decrease in the number of crashes of about 39% in the treated areas, confirming the advantage of the treatment. Furthermore, an improvement in skid resistance resulted in considerable safety benefits for both intersections and road segments (Lyon & Persaud, 2008). Pardillo Mayora and Jurado Piña (2009) stated that an improvement in skid resistance after resurfacing lead to significant decrease in crash rates as skid resistance increases in both wet and dry pavement conditions. They reported an average of 68% reduction in wet-pavement crash rate after resurfacing. Chassiakos, Panagolia, Theodorakopoulos, and Vagiotas (2004) found 44% reduction in crashes at sections related to rural, parts of urban and intersections due to resurfacing.

In addition, findings from previous studies indicate that an increase in pavement roughness results in an increase in crash rate (Chan, Huang, Yan, & Richards, 2010, 2009; Cairney & Bennett, 2008; Ihs, 2004; Othman, Thomson, & Lanner, 2009; Bester, 2003). Larson, Hoerner, Smith, and Wolters (2008) added that an increase in crash rates at signalised intersections is associated with an increase in pavement roughness due to the presence of rutting, shoving and corrugated road surface. This is in addition to an increase in the possibility of aggregate polishing as a result of regular stopping and starting operations of high traffic volume at signalised intersections. Dong, Clarke, Yan, Khattak, and Huang (2014) resulted in an increase in crash frequency due to increase in International Roughness Index (IRI) at 603 signalised intersections in Tennessee/USA through comparison of both multivariate random-parameters zero-inflated negative-binomial regression and multivariate zero-inflated negative binomial regression models. Al-Masaeid (1997) found that an increase in road roughness lead to reduction in single-vehicle crash rate and increase in multiple-vehicle crash rates.

Further, the relation of rutting to safety is a major concern; however, there is no clear relationship between rut depth and crashes (Ihs, Velin, & Wiklund, 2002). Deeper ruts are not correlated with higher crash risk except for night and rain weather-related crashes (Ihs, Gustafsson, Eriksson, & Wiklund, 2011). Cairney, Styles, and Bennett (2005) carried out a study on a rural major highway in Victoria to examine the relationship between rutting and crashes. The authors stated that pavement ruts of 20 mm and deeper lead to 60% increase in crash risk. Chan et al. (2009) indicated that rut depth did not affect crash prediction models significantly except for night and rain weather-related crashes. Furthermore, Start, Kim, and Berg (1998) found that crash rate at undivided rural highways increases with rut depths greater than 7.6 mm.

Many studies have been also concerned with the identification of geometric parameters that lead to crash occurrence. Different relationships were developed between geometric characteristics and crash occurrence at both signalised intersections and roadway segments. In Victoria, Australia, Ogden, Newstead, Ryan, and Gantzer (1994) carried out a study of a sample of 76 signalised intersections to identify factors affecting crash occurrence. Based on the analysis described in their study it was found that: (i) lower crash rates were observed in approaches with exclusive right-turn lanes, (ii) there was a tendency for approaches with narrow lanes (less than the typical width of 3–3.5 m) to have higher crash frequencies and (iii) the presence of medians greater than 0.9 m contributed to lower crash occurrence. Bonneson and Mccoy (1997) used negative binomial distribution to identify common trends of crash occurrence related to median treatment on urban arterials. Higher crash frequency was found for undivided cross sections (no median) in situations where parallel parking was allowed. Knuiman, Council, and Reinfurt (1993) found that total crashes and crash rates for particular types and severity decreased rapidly as the median width exceeded nearly 7.6 m.

Othman et al. (2009) reported that an increase in lane width leads to an increase in crash rates for all types of roads. This evidence is thought to be due to overtaking behavior, lane changing maneuvers and higher speeds on wide carriageways. Based on Ilion's country-data, Noland and Oh (2002) performed a study to evaluate the effect of infrastructure changes on crashes. The authors found that higher traffic related fatalities were associated with the increased number of lanes, lane width and outside shoulder widths. Jiang (2012) studied the effect of traffic engineering factors on the occurrence of two-vehicle crashes. The results showed that lane width did not produce any significant influence on crash severity and the effect of traffic engineering factors vary by the type of crashes. In contrast, Popoola, Abiola, and Odunfa (2018) found that the wider lanes are related to lower crashes on a two-lane highway.

A study to explore the effect of different factors on crash frequency and examine the effect of spatial correlation among 476 signalised intersections along 41 corridors in Florida was conducted by Abdel-Aty and Wang (2006). The authors found that approaches with a larger number of lanes were associated with higher crash frequency. Similarly, Milton and Mannering (1998) pointed out that the number of lanes in a section of highway was a significant factor in increasing crash frequency. Turner, Singh, and Nates (2012) evaluated 238 signalised intersections in New Zealand and Australia. The authors attempted to quantify the effect of various factors related to intersection geometry, signal phasing and land-use environment on different crash types and to develop crash prediction models. Results indicated that more crashes occurred at larger intersections with higher numbers of approach and exit lanes and at those with greater intersection depths (the crossing distance from one approach to the opposite approach) and the number of crashes rose at intersections with parking within 30–40 m of the stop line. The authors also found that presence of a shared lane had a mixed effect on crash occurrence. In peak period, right angle crashes increased in all cities with the presence of a shared lane. However, approaches with the presence of shared right turn/through lanes were associated with a reduction in crashes at intersections with lower traffic volumes. Furthermore, Wang and Abdel-Aty (2006) reported that with regard to type of right-turn lane (left-turn in Australia) on minor roadways, exclusive and channelised right-turn lanes were found to be related to fewer rear-end crashes than shared right turn lanes. However, in a study to investigate the safety performance of different right-turn designs at intersections, Fitzpatrick, Schneider, and Park (2006) found that shared through/right turn lanes were associated with lower crash frequency. Turner et al. (2012) stated that adequate phase time is a significant factor in reducing crash occurrence. Morgan, Tziotis, and Turner (2009) added that insufficient phase time is a possible contributory factor for different crash types.

The purpose of this study is to establish how these pavement condition parameters contribute to occurrence of casualty crashes at signalised intersections. Casualty crashes include all severity levels namely; fatal, serious injury and others. This is in addition to their interactions with situational factors including speed limit, traffic volume, light condition and surface moisture condition. The contribution of intersection geometric characteristics to safety performance after treatment is also assessed. The sites of signalised intersections used in this study were selected to have been subjected to surface treatment (thin asphalt surfacing) with no changes in geometric characteristics over the study period. Their safety performance was compared before and after treatment using suitable regression analyses. In both assessments contribution of human behaviour to crash occurrence is considered constant. Details of sites selection criteria, collection of surface condition and crash data, assessment approach are described in this paper. Assessment results are also presented and discussed.

2. Sites selection

The sites of signalised intersections used in this study were selected from a metropolitan region in Melbourne, Victoria/Australia. For assessing the effects of surface condition parameters and other situational factors on safety performance, only sites that satisfied all the following criteria were considered:

- Only sites that have been subjected to surface treatment are considered
- Have surface condition data (roughness, rutting and skid resistance) for the year before and year after treatment year
- Have crash data over 3–5 years before and after treatment year
- Did not have any changes in geometry or speed limit during the assessment period

At the time of this study, historical data for surface treatment of pavements in the selected region was available for the period of 2007–2010. All signalised intersections that were subjected to surface treatment during this period were identified but considering the remaining above criteria, a sample of 57 sites was selected. It is important to note that surface treatments applied to any of these intersections do not cover the whole intersection. They may cover the intersection centre, intersection centre and approaches or any of the approaches (immediate 200 m). The total length of treatment for sites that have been included in the analysis ranges between 100 and 500 m.

Intersections with tram tracks (light rail) were also excluded to control their effect on the different pavement condition parameters. Since no changes were made to the geometry or speed of these intersections post treatment, it is safe to assume that the only variable road-related factors are surface condition and the normal annual increase in traffic volumes.

A sample of these treated intersections with different geometric characteristics was selected. The geometric characteristics for the selected intersections were collected for the years after treatment, to control for the effect of pavement surface condition. Considering the significant contribution of phase time, sample size was limited to those with available phase time data for the years of interest. A total of 49 sites were selected and Google Earth was used to identify their geometric characteristics.

3. Statistical approaches

Crash data are non-negative integers and when the mean is small, their distributions will be positively skewed. Thus, Poisson/NB are appropriate primary models to consider for crash data (Hilbe, 2011). Numerous studies have been carried out to analyse crash occurrence at intersections by applying statistical approaches including Poisson and Negative Binomial (NB) regression models. One of the characteristics of crash frequency data is that the variance of crash counts is greater than the mean (over dispersion). The assumption of Poisson regression (equi-dispersion) is that the mean and variance of crash count data are equal. Milton and Mannering (1998) suggested that the negative binomial regression is an appropriate predictive model for applying in crash frequency studies.

In this study, the crash data are not independent owing to the repeated observations of the effect of light and surface moisture conditions over four occasions (day-dry, day-wet, night-dry and night-wet) for each intersection. This clustering needs to be taken account of in the statistical analysis to compute appropriate standard errors. To account for the repeated observations, the Generalised Estimating Equation (GEE) approach has been used to account for the correlation among clustered data (Liang & Zeger, 1986; Diggle, Heagerty, Liang, & Zeger, 2002). Accordingly, because the observations are clustered within subjects and are not time series data, an exchangeable correlation matrix with robust variance estimation was applied for crash severity analysis.

A GEE gives a marginal (or population average) interpretation of results which is identical to results obtained using independent data. Furthermore, for a GEE, the estimated effects are robust to misspecification of the correlation structure for reasonable samples sizes (Diggle et al., 2002). Alternatives to GEEs include random effects models, robust variance estimation or simple scaling of variances (Diggle et al., 2002; Hilbe, 2011). A disadvantage of random effects models is that they require the correlation structure to be correct for the estimated effects to be unbiased. Moreover, random effects models for discrete data (such as crash counts) give a "subject-specific" interpretation which is not necessarily comparable to results obtained in studies using independent data.

4. Data preparation

The data collected for studying safety performance of the identified signalised intersections include crash data, surface condition, geometric characteristics, speed limit, phase time and traffic volumes. Preparations of these data sets are described in the following subsections.

Crash data was collected from CrashStat database (VicRoads, 2014) which contains information on crashes that occurred during the 13-year period (2000–2013) of interest. This database was filtered to select only crashes that occurred at the treated direction of each selected site, using the chainages where crashes occurred. Crash location could be within centre of intersection or immediate 200 m of its approaches as shown in Fig. 1. Crash data for these intersections was available either for 3, 4 or 5 years before and after treatment years. To provide a balanced analysis, an equal period of crash data for before and after treatment was selected for each intersection. And to maximise the power of analysis the longest available period for each intersection (i.e. either 3, 4 or 5 years) was used. Crash data collected includes all casualty crashes i.e. covering all severity levels (fatal, serious injuries and other), type of crash (head on, rear etc.), light condition when they occurred (day or night), road surface moisture condition (wet or dry) and speed limit.

Surface condition data including roughness, rutting and skid resistance of the treated direction of each site were collected for one year before treatment year and one year after. Roughness and rutting data are collected for the whole network bi-annually using laser profiler. Roughness determined from longitudinal profile measurements is reported in terms of the International Roughness Index in m/km for 100 m segments. The measure used herein is lane IRI i.e. average IRI of both wheel paths. Rutting is determined from transverse profile measurements and reported in terms of average lane rut depth in mm for 100 m segments. Skid resistance data is collected regularly using Side-ways force Coefficient Routine Investigation Machine (SCRIM). It is collected for both wheel paths and reported for each separately and their average in terms of Side-ways force Coefficient (SFC) values, also for 100 m segments. Data documented by SCRIM is positive integer equivalent to the SFC*100.

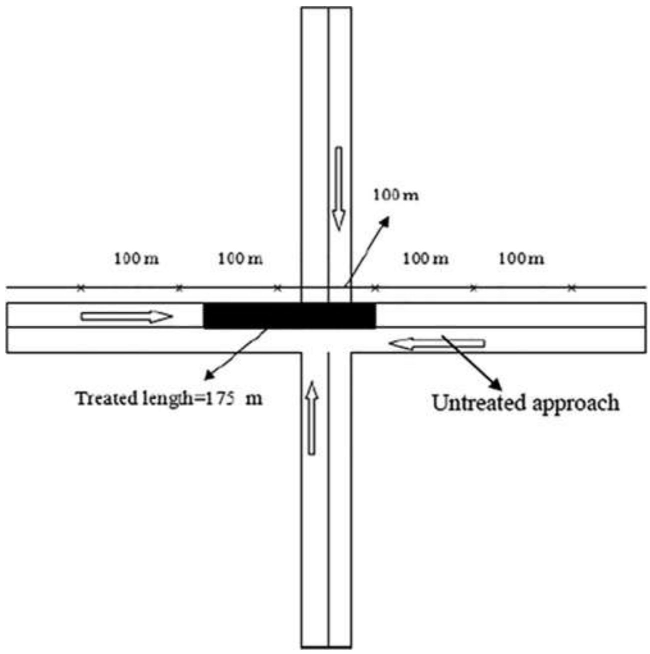


Fig. 1. Sketch of an intersection and approaches of immediate of 200 m illustrating the treated length.

For the analysis performed herein, data of each condition parameter for each site was averaged over a length of 500 m covering the intersection centre and a maximum of 400 m of its approaches. It is important to note that surface treatments of pavements at intersections and their approaches are currently triggered by surface distress ratings, referred to as Surface Inspection Rating (SIR), regardless of roughness level. SIR of an asphalt surfacing is a composite index of the ratings for cracking, stone loss, texture loss, patching and deformation (VicRoads, 2004).

Geometric characteristics and speed limit data have been collected for the treated approaches of the selected sample through site maps and special tools such as Google Earth. The geometric characteristics that have been collected are: approach width, number of approach lane, intersection depth, presence of median, presence of shared lane, presence of bus stop, presence of parking and type of intersection. Detailed explanation of these variables is shown in Fig. 2. The intersection depth is the distance travelled by an approaching vehicle to reach to the opposite approach (Turner et al., 2012). There were 36 four legged intersections and 13 three-legged intersection that used for geometrics characteristics assessment. When the treated approach was on the major leg of four-legged or three-legged intersections, the intersection depth was measured from the limit line of the approach leg to that of the exit leg. In the case of three-legged intersections, when the treated approach was on the minor leg of the intersection, the intersection depth was the distance from the limit line of the approach leg to the opposite edge of the intersection.

Data for phase time of each intersection approach during morning peak hours (7–9 am) and afternoon peak hours (4–6 pm), for the year after treatment year were obtained. A phase time is the summation of the displayed green time and inter-green time (yellow time plus all-red time) (Akçelik, 2006). A detailed description of cycle-by-cycle of the operational characteristics of signalised intersections during a 24-hour period can be obtained through SCATS (Sydney Coordinated Adaptive Traffic System).

For assessing safety performance of intersections before and after treatment, other explanatory variables have been considered together with pavement conditions parameters and their interactions. They include light condition (day/night), road surface moisture condition (wet/dry), traffic volume and speed limit. Light and surface moisture condition were obtained from relevant crash database. For each site, traffic volumes were collected for the 3–5 years before and after treatment i.e. covering the same number of years of available crash data. Traffic volume data used in the analysis is in terms of Annual Average Daily Traffic (AADT) that uses the section of road where the intersection site is located i.e. not peak traffic at the intersection.

5. Assessing the effect of pavement surface condition

Pavement surface condition parameters in terms of roughness, rutting and skid resistance (averaged over a length of 500 m covering the intersection and a maximum of 400 m of its approaches),

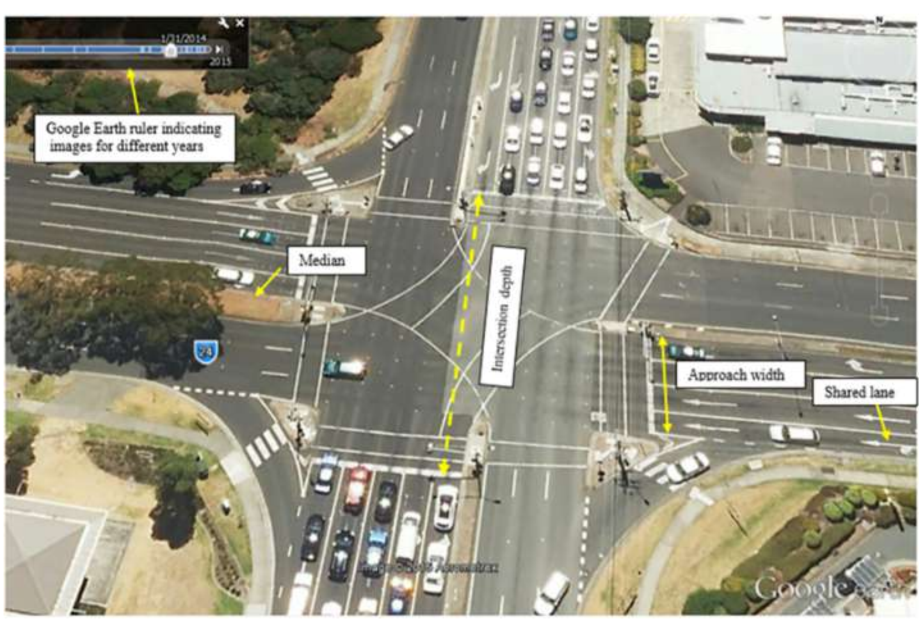


Fig. 2. Data collection of different geometric characteristics.

Table 1
Descriptive statistics of crash data and continuous variables for before and after treatment.

Variables	Min	Max	Mean	Std. Deviation	Interquartile range (75%–25%)
Before treatment					
Casualty crashes (3–5 years)	0	14	1.94	2.85	2–0
Roughness	1.55	5.25	3.03	0.83	3.6–2.41
Rutting	2.50	15.25	6.08	2.65	6.8–4.5
Skid resistance	0.39	0.70	0.51	0.07	0.55–0.47
AADT	1,600	27,136	13,955	5,853	17,786–10,000
Speed limit	60	80	71.23	6.59	80–70
After treatment					
Casualty crashes (3–5 years)	0	13	1.42	2.33	1–0
Roughness	1.81	4.78	2.88	0.59	3.26–2.48
Rutting	2.80	10.40	5.48	2.32	6.2–4
Skid resistance	0.48	0.71	0.60	0.06	0.64–0.58
AADT	1,600	28,550	14,689	6,183	19,000–10,000
Speed limit	60	80	71.23	6.59	80–70

traffic volume in terms of AADT, speed limit, light condition (day coded as 1 and night coded as 0) and surface moisture condition (dry coded as 1 and wet coded as 0) were explored for their effect on crash occurrence before and after resurfacing. Summary statistics of crash data and the continuous variables used in this study for before and after treatment are given in Table 1 for before and after treatment. Light and surface moisture conditions are categorical variables.

The assessment approach adopted herein to assess how pavement surface condition affects crash occurrence is to develop regression models linking each crash data set to corresponding variables and their interactions. Variables contributing to crash occurrence are then identified as those that have statistically significant effects in the models. In this study, several criteria were used to choose the appropriate probability model for crash data analysis. The following subsections present the testing performed to establish the best approach for modelling each data set.

5.1. Evaluation of over dispersion for casualty crashes

Numerous studies have been carried out to analyse crash occurrence at intersections by applying statistical approaches including Poisson and NB regression models and using Maximum Likelihood Estimation (MLE). One of the characteristics of crash frequency data is that the variance of crash counts is greater than the mean (over dispersion). The assumption of Poisson regression (equidispersion) is that the mean and variance of crash count data is equal. In the case of over-dispersion in crash frequency data, using a common Poisson model can lead to biased parameter estimates and erroneous conclusions concerning factors related to crash frequency data (Park & Lord, 2007). Milton and Mannering (1998) suggested that the negative binomial regression is an appropriate predictive model for applying in crash frequency studies.

In this study, several criteria were used to choose the appropriate probability model for crash data analysis. These criteria are: examining goodness of fit of the data to the probability mass function, examining whether the dispersion statistic is greater than one and testing whether the dispersion parameter > 0 using the Lagrange multiplier test.

Histograms of crash data sets for before and after treatment include high number of zero values and the variance of observed crash counts is greater than the mean (over dispersion). To choose the probability model for casualty crash data, goodness of fit was examined by plotting the observed crash data against both NB and Poisson Probability Mass Functions (PMFs) as given in Fig. 3. To construct the figure, first the mean and dispersion parameter for the Poisson and NB models were estimated. Then, these values

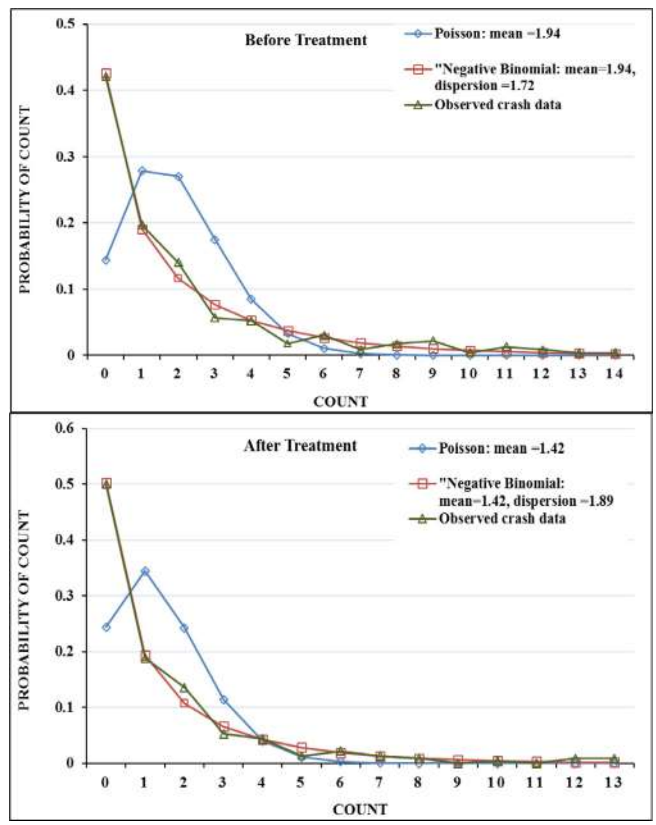


Fig. 3. Observed casualty crash frequency against Poisson and NB probability mass function, before and after treatment.

were substituted into the formulas for their respective PMFs to compute the count probabilities over the range of the observed data. Fig. 3 indicates that for both before and after treatment the observed number of casualty crashes fitted substantially better to the NB than the Poisson distribution.

The values of the dispersion statistic (Pearson chi square/degree of freedom) for the Poisson model with no covariates (intercept only) of 4.19 for before treatment and 3.84 for after treatment are greater than 1 which suggest that the dependent variable (crash frequency) is over dispersed. In addition, the dispersion parameter (alpha, α) refers to the parameter used in NB2 model to take account of over-dispersion which is assumed to be zero in the Poisson model. However, the negative binomial model allows it to be greater than zero. The dispersion parameters for

Table 2
GEE regression with NB and log link-before treatment.

Parameter	Coefficients β	Std. Error	P-Value	Exp (β)	95% Wald Confidence Interval for Exp (β)	
					Lower	Upper
Intercept	1.432	0.233	0.000	4.189	2.651	6.617
Light condition, Night = 0	−1.010	0.116	0.000	0.364	0.290	0.457
Light condition, Day = 1*	0	.	.	1	.	.
Surface MC, Wet = 0	−1.699	0.168	0.000	0.183	0.132	0.254
Surface MC, Dry = 1*	0	.	.	1	.	.
CRoughness	0.088	0.143	0.537	1.092	0.825	1.446
CRutting	−0.044	0.056	0.428	0.957	0.857	1.067
CSkid resistance × 100	−0.037	0.022	0.089	0.964	0.923	1.006
Speed Limit	0.021	0.183	0.907	1.022	0.714	1.462
CLogAADT	0.006	0.2	0.977	1.006	0.679	1.488
Light condition = 0 × Surface MC = 0	0.800	0.231	0.001	2.225	1.416	3.498
CSkid resistance × CLogAADT	−0.064	0.030	0.035	0.938	0.884	0.996
(Negative binomial)	0.593					
Number of observations	57*4 = 228					

*Reference case for category variable.

NB2 model with no covariates (intercept only) of 1.72 for before treatment and 1.89 for after treatment are greater than zero which again indicate that the dependent variable (casualty crash frequency) is over dispersed.

Another commonly used statistic to quantify the amount of over dispersion in crash data is Lagrange multiplier test. This test provides a p-value for which a decision can be made on whether to use Poisson or Negative Binomial model. Results of the Lagrange multiplier test for both before and after treatment crash data showed that the Z-score test is 5.328 for before treatment and 3.839 for after treatment with a t-probability of $P < 0.001$ for both. The significant Lagrange tests for both data sets indicates that the model dispersion parameter is different from zero. These results prove that the hypothesis of no over dispersion is rejected and the real over dispersion exists in both data sets. Therefore, casualty crash data, for before and after treatment, should be modeled as negative binomial as it is preferred over Poisson model.

5.2. Results and discussion

The crash data are not independent owing to the repeated observation of each intersection on four occasions (night-wet, night-dry, day-wet and day-dry). This clustering needs to be taken account of in the statistical analysis, e.g. to compute appropriate standard errors. The GEE method is an extension of the Generalised Linear Model (GLM) that accounts for the correlation among clustered data (Liang & Zeger, 1986).

Based on the goodness of fit evaluation of the model described in the previous section, this section applies the traditional NB regression as an initial step to estimate the dispersion parameter for the largest possible model including all predictors and significant interactions. GEE with NB distribution and log link function was used to analyse casualty crash data as the response variable. The explanatory variables including roughness, rutting and skid resistance were factors of interest in analysing casualty crash data. Natural log of traffic volume in terms of AADT (LogAADT), speed limit, light condition and surface moisture condition were included as control variables. This is done to focus the analysis and deal with interactions. With a small range of skid resistance from 0.39 to 0.7 for before treatment and 0.48 to 0.71 for after treatment (refer to Table 1), the effect of skid resistance × 100 was used to increase in the interpretability of regression coefficients.

Sometimes the effect of an independent variable, on a dependent variable, changes with the change in another variable which is called the interaction effect. To limit the number of interactions, all first order interactions (two-way interactions) among factors of interest and all first order interactions between each factors of interest and each control variable were examined. It was necessary to transform the continuous independent variables to centered variables by subtracting their mean values from their actual values. It is suggested that using centered variables in regression analysis leads to an increase in the interpretability of regression coefficients (Afshartous & Preston, 2011). The variables that were centered in this analysis are roughness, rutting, skid resistance and logAADT. Speed limit was not centered to the mean value, but to 60 km/h, which was the slowest speed limit.

For modelling GEE with NB distribution and log link function, a traditional negative binomial with all predictors and interactions was used to estimate the maximum likelihood value of (α) as a dispersion parameter. This value was inserted to GEE negative binomial model with all predictors and interactions.

By using GEE with the chosen correlation matrix (exchangeable), non-significant interactions were excluded and only predictors with main effect and significant interactions remained in the final model. Crash data was modelled using GEE for both before and after treatment. This working correlation matrix was based on a total of 57 clusters (number of signalised intersections). This resulted in a 4×4-dimension correlation matrix and the total number of observations = $57 \times 4 = 228$. The GEE was fitted using the value of $\alpha = 0.593$ for before treatment and $\alpha = 0.661$ for after treatment. Modelling outputs are presented next, separately, for before and after treatment casualty crashes.

1) Before treatment: The output results for before treatment is given in Table 2. The results show that not all predictors are statistically significant in explaining the variation in crash frequency. The factors with significant contributions are light condition, road surface moisture condition, their interaction product and interaction of skid resistance with traffic volume were the significant contributors to crash occurrence. The intercept is the expected number of crashes when all variables in the model are evaluated at zero. That is, during the day when it is dry, the expected total crash count is exp (1.432) = 4.2 crashes for an intersection with a speed limit of 60 km/h, if all other centered variables take their mean values. The main effect of light condition indicates that night time condition is associated with lower crashes as indicated by negative coefficient −1.01. In addition, the significant effect of surface moisture condition with negative coefficient of −1.699 implies that wet surface condition has lower crash frequency than dry surface condition.

To interpret the interaction effect between these two categorical variables, constructing a table is helpful. The predicted mean number of crashes for before period using GEE with NB model by

Table 3
Effect of light condition × surface moisture condition interaction on casualty crashes-before period.

	Predicted value of crash frequency	
	Wet	Dry
Night	0.62	1.53
Day	0.77	4.19

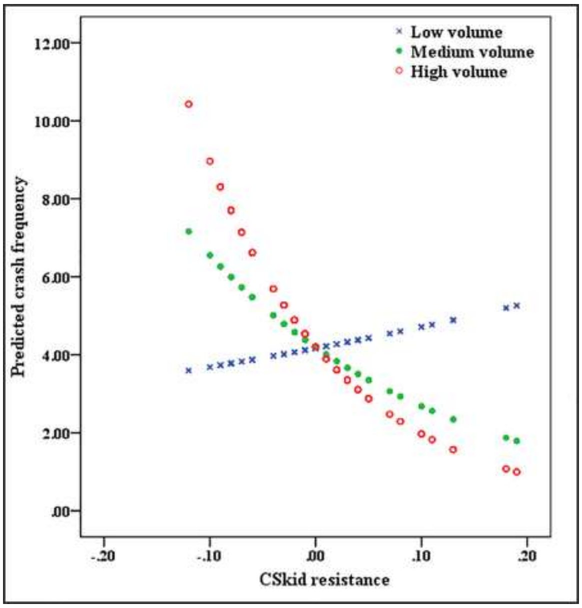


Fig. 4. Effect of skid resistance × traffic volume interaction on casualty crashes.

light and surface moisture condition setting all other centered variables to their mean values are given in Table 3.

The results indicate that crash frequency is highest in dry/day conditions followed by dry/night, wet/day with the lowest being wet/night conditions. To plot the effect of interaction between skid resistance and traffic volume on crash frequency, it is necessary to calculate the predicted value of crashes for before period as a function of skid resistance with representative values of traffic volume as shown in Fig. 4. Fig. 4 indicates that crash frequency increases with skid resistance for low traffic volume category. These findings could be attributed to the fact that both smoothness and skid resistance improvement following resurfacing may encourage drivers to drive faster hence leading to an increase in crash occurrence (Cleveland, 1987).

2) After treatment: The output results for after treatment is given in Table 4. Results illustrate that not all predictors are significant in crash occurrence. The intercept is the expected number of crashes when all variables in the model are evaluated at zero. That is, during the day when it is dry, the expected total crash count for after period is exp (1.31) = 3.7 crashes for an intersection with a speed limit of 60 km/h, if all other centered variables take their mean values. The main effect of light condition illustrates that night time condition is associated with lower crashes as indicated by the negative coefficient −1.457. In addition, the significant effect of surface moisture condition with negative coefficient of −1.869 implies that wet surface condition has lower crash frequency than dry surface condition.

For interpretation of interaction between these two categorical variables, the predicted mean number of crashes for after period, using GEE with NB model by light and surface moisture condition setting all other centered variables to their mean values are given in Table 5. The results indicate that the crash frequency increased in dry surface conditions and at day time as compared to wet surface conditions during night time. The effect of light condition on crash occurrence has changed by the effect of surface moisture condition.

The significant interaction term between roughness and light condition and roughness and surface moisture condition suggests that the relationship between crash frequency and roughness differs with light condition (day/night) and surface moisture condition (dry/wet). However, it is not entirely clear how it varied with day and night and with dry and wet surface condition. For interpretation of the significant interaction term between roughness and light condition the effect of this interaction was plotted as shown in Fig. 5.

Fig. 5 indicates that when the surface was rough, crash frequency was higher during day time than night time, because of higher traffic volume during the day as compared to night condition. Similar findings have been reported in Yan, Radwan, and Abdel-Aty (2005) and Clague and Baran (2005). However, the negative coefficient of the interaction term implies that at night time there is slightly less crash frequency at higher values of roughness. This finding is supported by a recent study by Buddhavarapu,

Table 4
GEE regression with NB and log link-after treatment.

Parameter	Coefficients β	Std. Error	*P*-Value	Exp (β)	95% Wald Confidence Interval for Exp (β)	
					Lower	Upper
(Intercept)	1.310	0.212	0.000	3.707	2.445	5.620
Light condition, Night = 0	−1.457	0.136	0.000	0.233	0.178	0.304
Light condition, Day = 1*	0	.	.	1	.	.
Surface MC, Wet = 0	−1.869	0.182	0.000	0.154	0.108	0.221
Surface MC, Dry = 1*	0	.	.	1	.	.
CRoughness	0.343	0.222	0.122	1.409	0.913	2.175
CRutting	−0.061	0.052	0.234	0.941	0.850	1.040
CSkid resistance × 100	−0.027	0.019	0.150	0.973	0.938	1.01
Speed Limit	−0.009	0.186	0.963	0.991	0.689	1.427
CLogAADT	−0.009	0.166	0.958	0.991	0.716	1.373
Light condition = 0 × Surface MC = 0	1.567	0.283	0.000	4.792	2.752	8.344
Light condition = 0 × CRoughness	−0.490	0.199	0.014	0.613	0.414	0.907
Surface MC = 0 × CRoughness	0.514	0.202	0.011	1.672	1.124	2.486
CRoughness × CLogAADT	0.762	0.288	0.008	2.142	1.217	3.770
(Negative binomial)	0.661					
Number of observations	57*4 = 228					

*Reference case for category variable.

Table 5
Effect of light condition × surface moisture condition interaction on casualty crashes-after period.

	Predicted value of crash frequency	
	Wet	Dry
Night	0.64	0.86
Day	0.57	3.71

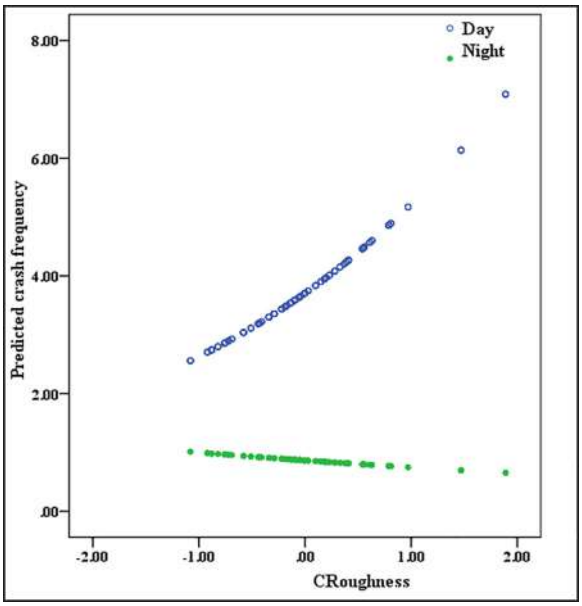


Fig. 5. Effect of roughness × light condition interaction on casualty crashes.

Banerjee, and Prozzi (2013), who related this to that drivers are more cautious while driving at night time conditions.

Similarly, for interpretation of the significant interaction term between roughness and surface moisture condition, the effect of this interaction was plotted as shown in Fig. 6. Fig. 6 indicates when the surface was rough crash frequency was higher at dry surface condition than wet surface condition. Crash frequency increased with increasing roughness at both wet and dry surface condition.

The results are consistent with previous studies. As pavement roughness increases, the crash rate increases which is related to the fact that the drivers are likely to lose control of their vehicles due to the presence of defective road conditions which affects the driver decision in changing driving speed suddenly (Al-Masaeid, 1997). Furthermore, as pavement condition deteriorates especially in wet weather condition, the crash rate increases (Chan et al., 2009). The probable explanation for this is that when the pavement surface is rough, the driver has more difficulty with the visibility of the road surface in adverse conditions which increases the likelihood of crash occurrence. The significant interaction term between roughness and traffic volume suggests that the relationship between crash frequency and roughness differs according to the levels of traffic volume as shown in Fig. 7. Fig. 7 indicates that crash frequency reduces with increasing roughness for low traffic volumes.

This finding may be explained by the reduction in operating speed as a result of reduction in the ride quality which is more evident at low volume intersections as drivers can easily drive at low speeds in intersections or change lanes to avoid defective road conditions. Similar findings have been reported by Al-Masaeid (1997) and Rodegerdts, Nevers, Robinson, Ringert, Koonce, Bansen, Nguyen, Mcgill, Stewart, and Suggett (2004). However, crash frequency increased as roughness increased for intersections with

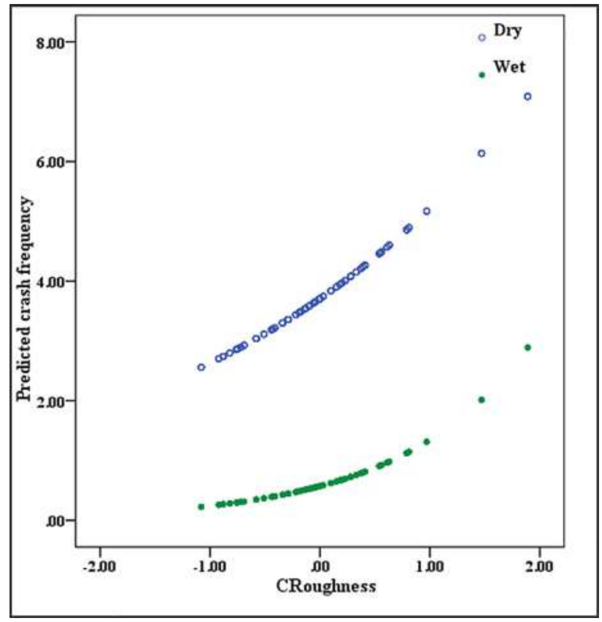


Fig. 6. Effect of roughness × surface moisture condition interaction on casualty crashes.

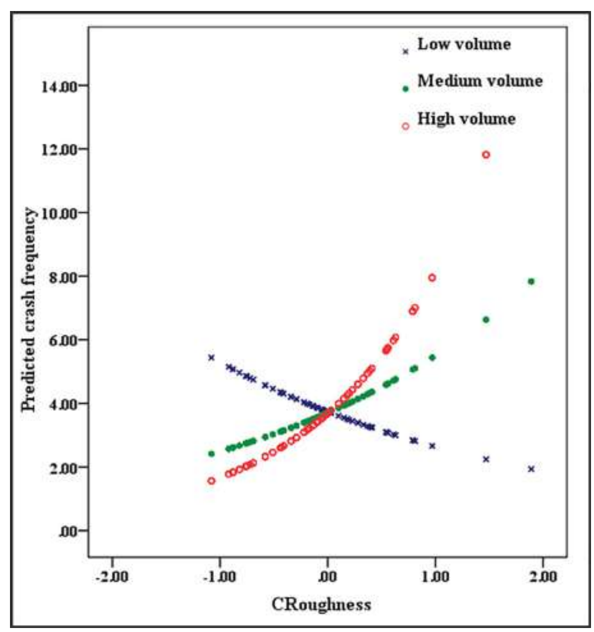


Fig. 7. Effect of roughness × traffic volume interaction on casualty crashes.

high traffic volumes which could be due to the presence of defective road condition and possible aggregate polishing. Similar results have been reported in Ihs (2004) and Larson et al. (2008).

6. Assessing the effect of geometric characteristics

The main purpose of this part of study was to assess the effects of geometric characteristics on safety performance of approaches of signalised intersections post resurfacing to control for the effect of pavement surface condition. The study aim was achieved through observing general trends and suitable statistical analysis. Descriptive analyses are performed to study the distribution of crash frequency with different geometric characteristics. A summary of observations from the analysis is provided in Table 6.

Table 6
Summary of observations from descriptive analysis.

Geometric characteristics	Summary of observations from descriptive analysis
Width of approaches	• Up to 50% of intersections in the sample have approaches 10–15 m wide • Crash frequency decreases as the approach width increases beyond 15 m
Intersection depth	• The depths of selected intersections range between 25 m and 49 m with about 63% of intersections have depths ranging between 30–40 m • Crash frequency increases at intersections with greater depths (i.e. >30 m)
Presence of median	• Medians are present at 88% of approaches in the sample set • Approaches with medians are associated with higher crash frequency than approaches with no medians
Presence of bus stop	• Bus stops are present at 33% of approaches in the sample set • Approaches with bus stops are associated with lower crash frequencies than approaches with no bus stops
Presence of shared lane	• Shared lanes are present at 49% of approaches in the sample set • Approaches with shared lanes are associated with higher crash frequency than approaches without shared lanes

Geometric characteristics including approach width, number of approach lanes, intersection depth, intersection type, presence of median, presence of shared lane, presence of bus stop, and presence of parking were initially considered to predict the number of crashes at signalised intersections after resurfacing. It is important to note that there were <10% of approaches with presence of parking in the sample set, so this factor was not considered in the statistical analysis. Among the highly correlated predictors, only the most significant variables were considered in the analysis. For example, because the number of approach lanes and approach width were found to be correlated, the former was dropped from the analysis. Also, intersection type was found to be correlated with phase time and therefore was dropped from the analysis.

In before and after assessment shown in previous section, surface roughness of treated approaches was found to have a significant effect on crash occurrence in the after treatment period, hence it is also considered in this assessment. In addition, traffic volume, speed limit, light condition (day coded as 1 and night coded as 0), surface moisture condition (dry coded as 1 and wet coded as 0) and phase time have been explored for their effects on total crash occurrences after resurfacing. Summary statistics of crash frequency and the continuous variables used in this part of study are given in Table 7. A related point to speed limit to consider, is that 28 observations are with speed limit of 60 km/h, 96 observations with speed limit of 70 km/h and 72 observations with speed limit of 80 km/h.

6.1. Evaluation of over dispersion in crash data post resurfacing

Similar criteria explained in Section 5.1 were used to choose the appropriate probability model for crash data. The histogram of crash data includes high number of zero values and a long positive skewness of the dependent variable (DV i.e. crash frequency). Furthermore, the variance of observed crash counts is greater than the mean suggesting over dispersion in a Poisson model. To choose the probability model for crash data, goodness of fit was examined by plotting the observed crash data against both NB and Poisson probability mass functions (PMFs) as given in Fig. 8. Fig. 8 indicates that the observed number of crashes fitted substantially better to the NB than the Poisson distribution.

The value of the dispersion statistic (Pearson chi square/degree of freedom) for the Poisson model with no covariates (intercept only) of 3.44 is greater than 1 which suggests that the dependent variable (crash frequency) is over dispersed. In addition, the dispersion parameter (alpha, α) refers to the parameter used in NB2 model to take account of over-dispersion. Poisson model assumes it to be zero, however, the negative binomial model allows it to be greater than zero. The dispersion parameter for NB2 model with no covariates (intercept only) of 1.9 is greater than zero which indicates that the dependent variable (crash frequency) is over dispersed. Results of the Lagrange multiplier test for crash data illustrate that the Z-score test is 3.568 with a t-probability of $P < 0.001$. A significant Lagrange indicates that the model dispersion parameter is different from 0. These results prove that the hypothesis of no over dispersion is rejected and that real over dispersion exists in the data set. Therefore, crash data should be modeled as negative binomial.

Based on the goodness of fit evaluation of the model described in the previous section, this section refers to the application of GEE with NB distribution and log link function used to assess the effect of geometric characteristics and other control factors. To limit the number of interactions, first order interactions (two-way interactions) among factors of interest and between factors of interest and each control variable were examined. To increase the interpretability of regression coefficients it was necessary to transform the continuous independent variables to centered variables by subtracting their mean values from their actual values. The variables that were centered in this analysis were roughness, phase time and logAADT. Speed limit was not centered to the mean value, but to 60 km/h, which was the lowest speed limit among the intersections studied.

To assess the effect of geometric characteristics and other control factors, GEE with NB distribution and log link function was used. The explanatory variables including approach width, presence of median, presence of shared lane, presence of bus stop, intersection depth and pavement roughness were factors of interest in analysing of crash frequency data. Traffic volume in terms of Annual Average Daily Traffic (AADT), speed limit, phase time, light condition and surface moisture condition were included as control variables. Crash data was modelled using GEE with working correlation matrix based on a total of 49 clusters (number of signalised intersections). This resulted in a 4×4-dimension correlation matrix and the total number of observations = $49 \times 4 = 196$. The GEE was fitted using the value of $\alpha = 0.321$. The detail output results using exchangeable correlation structure are given in Table 7.

Table 7
Descriptive statistics of crash data and continuous variables post resurfacing.

Variables	Minimum	Maximum	Mean	Std. Deviation	Interquartile range (75%–25%)
Crash frequency (4 years)	0	12	1.30	2.116	2–0
Roughness	1.77	3.86	2.8384	0.54355	3.21–2.46
AADT	1,600	28,438.00	15,983.20	5,696.81	19,321–10,000
Speed limit	60	80	71.23	6.59	80–70
Phase time	10.00	90.00	40.47	17.41	47–28

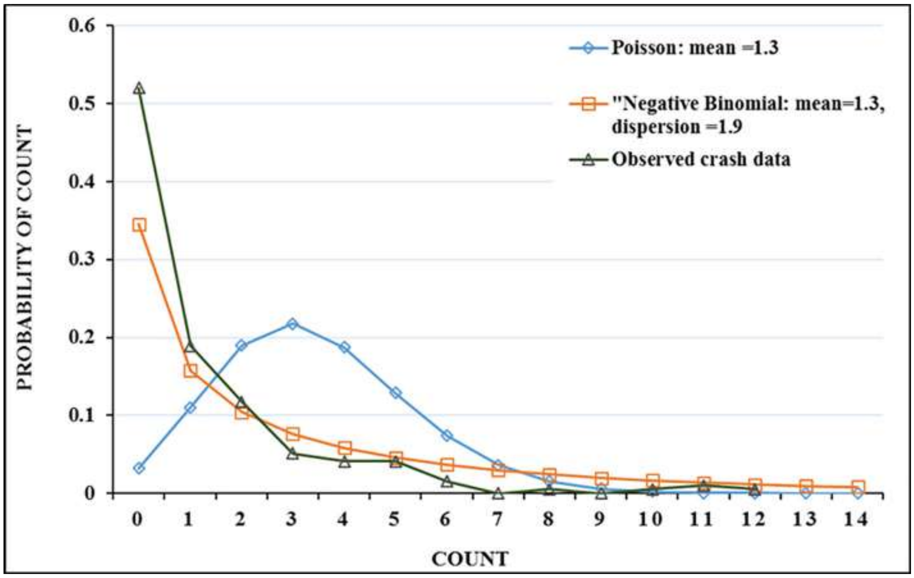


Fig. 8. Observed crash frequency against Poisson and NB probability mass function.

6.2. Results and discussion

This section interprets the results of analysing the effect of geometric characteristics and other control factors on crash data. It is important to note that the models include the significant and non-significant independent variables with main effect and only significant interaction products between explanatory variables. A summary of results from Table 8 is provided below.

1. Results in Table 8 illustrate that, intersection depth, roughness, speed limit, traffic volume and phase time are not significant in crash occurrence. However, light condition (day vs. night), surface moisture condition (dry vs. wet) and approach width in interactions with presence of shared lane, bus stop and median are significant in crash occurrence with p values of <0.05. It is important to note that as the overall p-value of the interaction product between approach width and presence of shared lane,

Table 8
GEE regression with NB and log link for crash data post resurfacing.

Parameter	Coefficients β	Std. Error	*P*-Value	Exp (β)	95% Wald Confidence Interval for Exp (β)	
					Lower	Upper
Intercept	0.749	0.555	0.178	2.114	0.712	6.276
Approach width < 12 m = 0	−1.858	1.392	0.766	0.156	0.010	2.388
Approach width 12 m–15 m = 1	1.690	0.607		5.420	1.650	17.804
Approach width ≥ 15 m = 2*	0	.		1	.	.
Shared lane, not present = 0	0.660	0.552	0.322	1.934	0.656	5.703
Shared lane, present = 1*	0	.		1	.	.
Bus stop, not present = 0	0.028	0.379	0.855	1.028	0.488	2.165
Bus stop, present = 1*	0	.		1	.	.
Median, not present = 0	0.127	0.734	0.170	1.135	0.270	4.779
Median, present = 1*	0	.		1	.	.
Intersection depth < 35 m = 0	−0.703	0.417	0.209	0.495	0.219	1.120
Intersection depth 35 m–40 m = 1	−0.606	0.405		0.546	0.247	1.207
Intersection depth ≥ 40 m = 2*	0	.		1	.	.
Light condition, Night = 0	−0.980	0.172	0.000	0.375	0.268	0.526
Light condition, Day = 1*	0	.		1	.	.
Surface MC, Wet = 0	−1.514	0.158	0.000	0.220	0.162	0.300
Surface MC, Dry = 1*	0	.		1	.	.
CRoughness	0.007	0.3128	0.982	1.007	0.546	1.859
Speed limit	−0.014	0.234	0.951	0.986	0.623	1.560
CLogAADT	0.848	0.616	0.169	2.334	0.698	7.805
CPhase time	0.005	0.007	0.481	1.005	0.992	1.018
Approach width = 0*Shared lane = 0	0.925	1.102	0.048	2.522	0.291	21.875
Approach width = 1.0* Shared lane = 0	−1.396	0.635		0.248	0.071	0.860
Approach width = 0* Bus stop = 0	0.921	0.724	0.019	2.511	0.607	10.382
Approach width = 1.0 * Bus stop = 0	−1.137	0.511		0.321	0.118	0.873
Approach width = 0 * Median = 0	2.260	0.925	0.015	9.583	1.564	58.720
(Negative binomial)	0.321					
Number of intersections (number of clusters)	49					
Size of cluster	4					
Number of observations	49*4 = 196					

bus stop and median are statistically significant with p value of 0.048, 0.019 and 0.015 respectively, the interpretation of individual categories is reasonable. The correlation estimated by exchangeable structure is 0.142. This value indicates that the correlation among repeated observations should be accounted for.

2. The intercept is the expected number of crashes when all variables in the model are evaluated at zero. That is, at an approach with speed limit of 60 km/h, a width greater than 15 m, a shared lane, bus stop and median, during a dry day and an intersection depth greater than 40 m, the expected total crash count is exp (0.749) = 2.11 crashes per 4-year (the total over 4 years is used), if all other centered variables take their mean values.
3. The main effect of light condition indicates that night time is associated with lower crashes as indicated by the negative coefficient of −0.98. In addition, the significant effect of surface moisture condition with negative coefficient of −1.541 implies that wet surface has lower crash frequency than dry surface. These findings are more likely to be capturing exposure owing to higher number of vehicles entering intersections during day time and dry weather conditions.
4. To interpret the interaction effect between approach width and presence of a shared lane, constructing a table is helpful. The predicted number of total crashes over four years using the GEE with NB model by approach width and presence of shared lane, bus stop and median setting all other centered variables to their mean values are given in Table 9.

- **Interaction between approach width and presence of shared lane:** The pattern of predicted crashes shown in Table 9 indicates that generally wider approaches tend to have higher crash frequency than narrower approaches. The crash frequency is higher at approaches having widths between 12 m and 15 m and with the presence of a shared lane as compared to others. The pattern of the numbers in this table shows a clear interaction between approach width and presence of a shared lane. The effect of presence of shared lane varies for different categories of approach width. For approaches without shared lane, there is an increase in the predicted number of crashes from approaches with width < 12 m to approaches with widths between 12 m and 15 m. However, it reduces slightly for approaches with widths ≥ 15 m. This is because the middle category (approach width between 12 m and 15 m) exists in a major proportion of intersections that carry high traffic volumes.
- This pattern is not the exact case for the other category of shared lane. With the presence of shared lanes and at narrow approaches (width < 12 m), lower crash number is obtained. This is related to the small proportion of such intersections that have shared lanes in addition that they carry low traffic volumes. Furthermore, the results illustrate that predicted number of crashes increases substantially when approach width increases from 12 to 15 m and decreases substantially when approach width increases ≥15 m. Similar results were reported by Turner et al. (2012) who stated that presence of shared lane had mixed effect on crash occurrence. The effect was positive in peak hours at intersections with higher traffic volume and negative at low volume intersections.
- **Interaction between approach width and presence of bus stop:** The pattern of predicted crashes shown in Table 9 is similar to that of the interaction effect between approach width and presence of shared lane. The high crash frequency for approach width 12–15 m is due to the fact that presence of bus stops can decrease the effective width of the approach. The on-street bus stop at intersection approaches causes vehicles following the bus to stop completely or change lane when realise that the bus is stopping. This stop and go condition and changing lane affect the moving traffic and results in an increase in crash occurrence. These findings are consistent with previous studies (Chimba, Sando, & Kwigizile, 2010; Chin & Quddus, 2003). Conversely, there is a big reduction in the number of crashes at wider approaches (approach width ≥ 15 m). This is owing to the fact that when the bus stop is located at upstream of intersection (the case for intersections considered in this study) and in the course of red light, the queue behind the intersection becomes an obstacle to the movement of stopping bus that is waiting upstream the stop line. Therefore, longer queue will develop by the stopping bus and more vehicles accumulate upstream the bus stop. The stop and go condition of the following vehicles and changing lanes results in conflicts between vehicles. Such behaviour has minor effect on the crash occurrence at wider approaches as compared to narrower approaches. Bauer and Harwood (1996), Brewer, Bonneson, and Zimmerman (2002) and Yan (2009) stated that crash frequency decreases with increasing approach width. Brewer et al. (2002) concluded that fewer red light running is likely to occur at wider intersections. Similar results are reported in Yan (2009) in that drivers are more careful in running a red light at wider approaches.
- **Interaction between approach width and presence of median:** The pattern of predicted crashes shown in Table 9 indicates that approaches having width < 12 m with no median are associated with high crash occurrence and there is a slight reduction in the predicted number of crashes when approach width increases from 12 to 15 m. Then it reduces slightly with wider approaches (width ≥ 15 m). Furthermore, the results illustrate that the presence of median at narrow approaches contributes to lower crash frequency as compared to others. Similar results are reported in Wang and Abdel-Aty (2006). However, the reverse of this trend is predicted for approaches having widths of 12 m–15 m which is consistent with findings of Chin and Quddus (2003). Again, there is a big reduction in the predicted number of crashes when approach width increases ≥15 m.

Table 9
Effect of approach width × presence of shared lane, bus stop and median interaction on total crashes over 4-years.

Variable categories	Predicted value of crash frequency over 4-years					
	Shared lane not present = 0	Shared lane present = 1*	Bus stop not present = 0	Bus stop present = 1*	Median not present = 0	Median present = 1*
Approach width < 12 m = 0	1.61	0.33	0.85	0.33	3.5	0.33
Approach width 12 m–15 m = 1	5.49	11.46	3.78	11.46	2.57	11.46
Approach width ≥ 15 m = 2*	4.09	2.11	2.17	2.11	2.4	2.11

*Reference case for category variable.

7. Summary and conclusion

Pavement surface condition parameters, geometric characteristics and other factors that influence casualty crashes at signalised intersections have been identified through analysing crash data by GEE with NB and log link function. The key findings from these analyses are summarised below:

1. For casualty crash analysis, GEE with NB regression and log link function was the appropriate model in predicting crash frequency for both before and after treatment.
2. Generally, the results indicate that crash frequency increased in dry surface conditions and during day time as compared to wet surface conditions during night time. This finding was consistent for both before and after treatment.
3. For before treatment, crash frequency was higher at higher values of skid resistance for sites with low traffic volumes, However, there was a clear reduction in crash frequency as skid resistance increased for sites with high traffic volumes.
4. For after treatment, crash frequency increased with increasing roughness during day time, however, at night time there was slightly lower crash frequency at higher values of roughness. Further, crash frequency increased with increasing roughness in both wet and dry surface condition. Additionally, crash frequency reduced with increasing roughness at intersections with low traffic volumes but increased at intersections with high traffic volumes.
5. Crash occurrence is not solely dependent on any of the geometric elements considered in this study, rather it is dependent on the interactions between the variables considered. It was found that the effects of presence of shared lane, bus stop or median on crash frequency vary for different categories of approach width.

Overall the findings of this study shed considerable light on the effects of pavement surface condition on safety performance of signalised intersections. Pavement surface condition parameters, roughness and rutting, need to be considered in either intervention criteria or treatment selection to ensure that they are reduced to acceptable levels for users' comfort and safety. The results also confirm the importance of considering interaction effects in crash analysis. The approach used in selecting the sites and analyses approaches used herein proved to be successful in addressing the study objectives with reasonable accuracy. The results reported herein highlight that when modelling crash data, both dispersion and correlation among repeated observations should not be ignored due to biased parameters that lead to erroneous estimates.

Conflict of interest

There is no conflict of interest, and the road authorities that provide all the data and information are happy to share and support this paper.

Acknowledgement

The authors express their appreciations to the different road agencies of Victoria for providing all data and relevant information for this study.

References

Abdel-Aty, M., & Wang, X. (2006). Crash estimation at signalized intersections along corridors: Analyzing spatial effect and identifying significant factors. *Transportation Research Record: Journal of the Transportation Research Board, 1953*, 98–111.
Akçelik, R., 2006. SIDRA intersection user guide. Akçelik, R. and Associates Pty, Greythorn, Victoria, Australia.
Afshartous, D., & Preston, R. A. (2011). Key results of interaction models with centering. *Journal of Statistics Education, 19*, 1–24.
Al-Masaeid, H. R. (1997). Impact of pavement condition on rural road accidents. *Canadian Journal of Civil Engineering, 24*, 523–531.
Bauer, K., Harwood, D. W., 1996. Statistical models of at-grade intersection accidents, FHWA-RD-99-094.
Bester, C. J., 2003. The effect of road roughness on safety. Presented at 82nd annual meeting. Washington DC: Transportation Research Board.
Bonneson, J. A., & Mccoy, P. T. (1997). Effect of median treatment on urban arterial safety an accident prediction model. *Transportation Research Record: Journal of the Transportation Research Board, 1581*, 27–36.
Brewer, M., Bonneson, J., Zimmerman, K., 2002. Engineering countermeasures to red-light-running. Today's transportation challenge: Meeting our customer's expectations. Washington, DC.
Buddhavarapu, P., Banerjee, A., & Prozzi, J. A. (2013). Influence of pavement condition on horizontal curve safety. *Accident Analysis & Prevention, 52*, 9–18.
Cairney, P., Styles, E., Bennett, P., 2005. Relationship between rutting, other road characteristics and crashes. ARRB Group, contract report VC 71038-1.
Cairney, P., Bennett, P., 2008. Relationship between road surface characteristics and crashes on Victorian rural roads. ARRB conference, 23rd, Adelaide, South Australia.30 July-1 August 2008.
Candappa, N., Scully, J., Newstead, S., Corben, B., 2007. Findings on the effectiveness of intersection treatments included in the Victorian State wide accident black spot program. Melbourne. Australasian Conference on Road Safety. Monash University Accident Research Centre.
Chan, C. Y., Huang, B., Yan, X., & Richards, S. (2010). Investigating effects of asphalt pavement conditions on traffic accidents in Tennessee based on the pavement management system (PMS). *Journal of Advanced Transportation, 44*, 150–161.
Chan, C. Y., Huang, B., Yan, X., & Richards, S. (2009). Effects of asphalt pavement conditions on traffic accidents in Tennessee utilizing pavement management system. *88th Annual Meeting*. Washington DC: Transportation Research Board.
Chassiakos, A., Panagolia, C., Theodorakopoulos, D., Vagiotas. P., 2004. Evaluating the effectiveness of pavement improvements in accident reduction. 6th international conference on managing pavements: The lessons, the challenges, the way ahead, 19–24 October 2004, Brisbane, Queensland, Australia.
Chimba, D., Sando, T., & Kwigizile, V. (2010). Effect of bus size and operation to crash occurrences. *Accident Analysis & Prevention, 42*, 2063–2067.
Chin, H. C., & Quddus, M. A. (2003). Applying the random effect negative binomial model to examine traffic accident occurrence at signalised intersections. *Accident Analysis and Prevention, 35*, 253–260.
Clague, S., Baran, E., 2005. Skid resistance: District implications and applications. Road system and engineering technology forum, 2005, Brisnane, Queensland, Australia.
Cleveland, D. (1987). *Effect of resurfacing on highway safety. State of the art report 6. United States of America: A synthesis of prior research*. Transportation Research Board, National Research Council.
Diggle, P. J., Heagerty, P., Liang, K. L., & Zeger, S. L. (2002). *Analysis of longitudinal data* (2nd ed). Oxford: Clarendon Press.
Dong, C., Clarke, D. B., Yan, X., Khattak, A., & Huang, B. (2014). Multivariate random-parameters zero-inflated negative binomial regression model: An application to estimate crash frequencies at intersections. *Accident Analysis & Prevention, 70*, 320–329.
Fitzpatrick, K., Schneider, W. H., & Park, E. (2006). Operation and safety of right-turn lane designs. *Transportation Research Record: Journal of the Transportation Research Board, 1961*, 55–64.
Giles, C. G., & Sabey, B. E. (1959). Skidding as a factor in accidents on the roads of Great Britain. In *Proceedings, First International Skid Prevention Conference* (pp. 27–40).
Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge, UK: Cambridge University Press.
Ihs, A., 2004. The influence of road surface condition on traffic safety and ride comfort. 6th international conference on managing pavements: The lessons, the challenges, the way ahead. 19–24 October 2004 Brisbane Queensland, Australia.
Ihs, A., Gustafsson, M., Eriksson, O., Wiklund, M., 2011. Road user effect models: The influence of rut depth on traffic safety, VTI rapport 731A. Prepared for the Finnish Transport Agency.
Ihs, A., Velin, H., & Wiklund, M. (2002). *The influence of road surface condition on traffic safety: Data from 1992–1998*. VTI Meddelanden.
Jiang, X. (2012). *Effects of highway geometric features and pavement quality on traffic safety* PhD dissertation. Knoxville: The University of Tennessee.
Kinnear, D., Lainson, L., Penn, H., 1984. An evaluation of four pavement treatments in sydney in relation to crash reduction, skid resistance and cost. 12th Australian Road Research Board Conference, 27-31 August 1984 Hobart, Tasmania, Australian Road Research, 12(3), 1-17.
Knuiman, M. W., Council, F. M., & Reinfurt, D. W. (1993). Association of median width and highway accident rates. *Transportation Research Record*: *Journal of the Transportation Research Board, 1401*, 70.
Larson, R. M., Hoerner, T. E., Smith, K. D., Wolters, A. S., 2008. Relationship between skid resistance numbers measured with ribbed and smooth tire and wet accident locations. FHWA/OH-2008/11, USA, State Job Number 134323.
Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika, 73*, 13–22.

Lyon, C., & Persaud, B. (2008). Safety effects of targeted program to improve skid resistance. *Transportation Research Record: Journal of the Transportation Research Board, 2068*, 135–140.

Milton, J., & Mannering, F. (1998). The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies. *Transportation, 25*, 395–413.

Morgan, R., Tziotis, M., Turner, B., 2009. Guide to road safety–part 8: Treatment of crash locations. AGRS08/09, Austroads, Sydney, Australia.

Noland, R.B., Oh, L., 2002. The effect of infrastructure and demographic change on traffic related fatalities and crashes: A case study of Illinois county-level data. 77th Annual Meeting of the Transportation Research Board. Washinton, DC.

Noyce, D. A., Bahia, H., Yambo, J. M., & Kim, G. (2005). *Incorporating road safety into pavement management: Maximizing asphalt pavement surface friction for road safety improvements*. Midwest Regional University Transportation Center Traffic Operations and Safety (TOPS) Laboratory.

Ogden, K. W., Newstead, S., Ryan, P., & Gantzer, S. (1994). *Factors affecting crashes at signalised intersections*. Melbourne, Australia: Monash University Accident Research Centre.

Oliver, J. (1999). *Review of skid resistance and accident relationship* (p. 91153). Vermont South Victoria, RC: ARRB Transport Research.

Othman, S., Thomson, R., Lanner, G., 2009. Identifying critical road geometry parameters affecting crash rate and crash type. Annual proceedings/association for the advancement of automotive medicine, 2009. Association for the Advancement of Automotive Medicine 53, 155-165.

Pardillo Mayora, J. M., & Jurado Piña, R. (2009). An assessment of the skid resistance effect on traffic safety under wet-pavement conditions. *Accident Analysis & Prevention, 41*, 881–886.

Park, E. S., & Lord, D. (2007). Multivariate Poisson-lognormal models for jointly modeling crash frequency by severity. *Transportation Research Record: Journal of the Transportation Research Board, 2019*, 1–6.

Popoola, M. O., Abiola, O. S., & Odunfa, S. O. (2018). Effect of traffic and geometric characteristics of rural two lane roads on traffic safety: A case study of Ilesha-Akure-Owo road, South-West, Nigeria. *Journal of Engineering and Technology, 3*, 2579–10617.

Rodegerdts, L. A., Nevers, B., Robinson, B., Ringert, J., Koonce, P., Bansen, J., Nguyen, T., Mcgill, J., Stewart, D., Suggett, J., 2004. Signalised intersections: Informational guide.

Saplioglu, M., Eriskin, E., Yuzer, E., & Aktas, B. (2012). Investigation skid resistance effects on traffic safety at urban intersections. *10th International Congress on Advances in Civil Engineering. 17-19 October 2012 Ankara, Turkey*.

Simpson, C., Eng, B., 2005. Performance of high skid resistant surfaces-crash trends, Australia.

Start, M., Kim, J., & Berg, W. (1998). Potential safety cost-effectiveness of treating rutted pavements. *Transportation Research Record: Journal of the Transportation Research Board, 1629*, 208–213.

Turner, S., Singh, R., & Nates, G. (2012). Crash prediction models for signalised intersections: Signal phasing and geometry. *NZ Transport Agency research report 483, Christchurch, New Zealand*.

VicRoads (2004). *Surface inspection rating procedure for pavements surfaced with sprayed seals and asphalt*. Melbourne/Australia: VicRoads.

VicRoads, CrashStats, 2014 [online]. Available from: http://www.vicroads.vic.gov.au/Home/SafetyAndRules/AboutRoadSafety/StatisticsAndResearch/CrashStats.htm. [Accessed January 1, 2014].

Wang, X., & Abdel-Aty, M. (2006). Temporal and spatial analyses of rear-end crashes at signalized intersections. *Accident Analysis & Prevention, 38*, 1137–1150.

Yan, F., 2009. Association of geometric characteristics at intersections with red light running frequencies. Masters Thesis, 2009.

Yan, X., Radwan, E., & Abdel-Aty, M. (2005). Characteristics of rear-end accidents at signalised intersections using multiple logistic regression model. *Accident Analysis and Prevention, 37*, 983–995.

Nasreen is a lecturer at university of Duhok-Iraq, Civil Engineering Department. She received her PhD from Swinburne University of Technology-Australia. Her research consists of assessing the impacts of pavement surface condition and geometric characteristics on safety and capacity performance of signalized intersections. Her research interests are transportation and traffic engineering. She is also interested in different statistical analysis techniques (Linear regression, Negative Binomial and Poisson Regression using Generalized Estimating Equation, Logistic regression and Empirical Bayes approach).

Rayya worked at Swinburne University of Technology (SUT), Civil Engineering, from 1998 to 2005. From 2005 to 2008, she worked at ARRB Group Ltd as a senior then as a Principal Research Engineer in the area of Asset Management and was responsible for managing a number of Austroads research projects. In 2008 she rejoined SUT. Research interests include pavement performance modelling and sustainability performance measurement of road asset management practices. Other research areas include using life cycle costing analysis in assessing the economic impacts of high productivity vehicles on road pavements, soil stabilisation and asphalt mix design.

Michael started with Swinburne as a biostatistician and epidemiologist in October, 2015. Previously, he had worked for statistics departments at Monash University, CSIRO and the University of Cambridge, where he did his PhD in biostatistics. His training and interest in epidemiology began when he did a Master of Public Health at the University of Sydney. He learnt most of what he know about study design and measurement methods by training in epidemiology. In 2020, he supervises a research fellow, co-supervise a PhD student and a Masters of Applied Statistics student on an industry supported project with Epworth Hospital.

Journal of Safety Research 76 (2021) 90–100

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Evaluating the speed camera sites selection criteria in the UK

Haojie Li [a,b,c,*], Manman Zhu [a,b,c], Daniel J. Graham [d], Gang Ren [a,b,c]

[a] School of Transportation, Southeast University, China
[b] Jiangsu Key Laboratory of Urban ITS, China
[c] Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, China
[d] Transport Strategy Centre, Imperial College London, UK

ARTICLE INFO

Article history:
Received 2 June 2020
Received in revised form 17 July 2020
Accepted 30 November 2020
Available online 17 December 2020

Keywords:
Speed Cameras
Selection Criteria
Propensity Score Matching
Empirical Bayes

ABSTRACT

Introduction: Speed cameras have been implemented to improve road safety over recent decades in the UK. Although the safety impacts of the speed camera have been estimated thoroughly, the criteria for selecting camera sites have rarely been studied. This paper evaluates the current speed camera sites selection criteria in the UK based on safety performance. *Method:* A total of 332 speed cameras and 2,513 control sites with road traffic accident data are observed from 2002 to 2010. Propensity score matching method and empirical Bayes method are employed and compared to estimate the safety effects of speed cameras under different scenarios. *Results:* First, the main characteristics of speed cameras meeting and not meeting the selection criteria are identified. The results indicate that the proximity to school zones and residential neighborhoods, as well as population density, are the main considerations when selecting speed camera sites. Then the official criteria used for selecting camera sites are evaluated, including site length (a stretch of road that has a fixed speed camera or has had one in the past), previous accident history, and risk value (a numerical scale of the risk level). The results suggest that a site length of 500 m should be used to achieve the optimum safety effects of speed cameras. Furthermore, speed cameras are most effective in reducing crashes when the requirement of minimum number of historical killed and seriously injured collisions (KSIs) is met. In terms of the risk value, it is found that the speed cameras can obtain optimal effectiveness with a risk value greater than or equal to 30, rather than the recommended risk value of 22.

1. Introduction

About 1.35 million road traffic deaths occur every year globally, placing a great burden on individuals, property, and society (World Health Organization, 2018). A number of safety measures have been proposed to improve road safety, such as traffic calming devices, road signs, and speed limit enforcement cameras (e.g., Li & Graham, 2016; Montella et al., 2015; Oviedo-Trespalacios et al., 2019). A critical issue when implementing these safety measures concerns which road sections or sites should be selected. As discussed in previous studies, the identification and selection of high-risk locations for highway safety management are mostly based on crash frequency and severity (e.g., Cheng & Washington, 2005; Montella, 2010; Manepalli & Bham, 2016). For example, specific rules for proposing speed camera sites are described in the handbook produced by the UK's Department for Transport (DfT, 2005). Speed camera sites selection criteria in the UK mainly depend on the number and severity of collisions. In addition, as reported by Transportation Research Board (NCHRP Report 729, 2012), selecting sites by the potential to reduce crashes, both for speed and red light cameras deployment, can be accomplished by reviewing crash frequency and crash rate.

Previous studies usually involve the evaluation of the safety effect of speed cameras, however, it remains unknown how much site selection criteria actually influence: (a) subsequent compliance with the specific criteria and (b) effectiveness of the implementation of speed cameras. Currently in the UK, site selection criteria for fixed speed camera sites exist (Gains et al., 2005; DfT, 2005) as shown: (1) number of killed and serious collisions (KSIs): at least 3 KSIs per km in the baseline period; (2) risk value required: at least 22 per km; (3) site length: between 400–1500 m; (4) 85th percentile speed at collision hot spots: 85th percentile speed at least 10% above speed limit; (5) percentage over the speed limit: at least 20% of drivers are exceeding the speed limit. In fact, the criteria for selecting camera sites was published in 2004 and improved in 2005 (DfT, 2004, 2005). It is still unclear

* Corresponding author at: School of Transportation, Southeast University, Jiu Long Hu, Nanjing 210000, China.
E-mail address: h.li@seu.edu.cn (H. Li).

https://doi.org/10.1016/j.jsr.2020.11.013
0022-4375/

how these rules are implemented in practice and whether the current criteria can achieve optimum safety benefits. For example, according to the handbook published in 2004, one of the criteria was at least 4 KSIs per km in the baseline period, while the number of KSIs was at least 3 in the 2005 version. More sites have been selected into the speed camera program due to the change of selection criteria. However, why the selection criteria were changed has not been discussed in the handbook. In addition, it is often the case that the criteria are not strictly met in practice, and sites not meeting the criteria may still be selected as exceptional sites for other safety concerns. Thus, it is also interesting to investigate the characteristics and safety performance of those sites that do not meet the stated criteria.

This study aims to evaluate the current UK speed camera sites selection criterion by analyzing the safety effects. In addition, this study investigates the variation in treatment effects due to the differences in site characteristics, specifically the extent to which the sites meet the selection criteria. The evaluation results can help to improve the safety performance of the current speed camera program, and hence will make the travel environment safer. In addition, this study will contribute to policy-making and legislation by providing new evidences and practical recommendations on enhancing the effectiveness of speed camera assignment. Moreover, the output from this research may be helpful to the improvements in economic performance of speed cameras. However, the cost-effectiveness of current speed camera deployment strategies cannot be evaluated due to the data restriction.

The propensity score matching (PSM) method and empirical Bayes (EB) method are both used in this study. The PSM method can control for the regression to the mean (RTM) and confounding factors, which cannot be fully addressed by the conventional before-after or cross-sectional studies (Rosenbaum & Rubin, 1983; Sasidharan & Donnell, 2013). RTM is a statistical phenomenon that occurs in before-after study designs that target "high risk" groups (e.g., sites with a high number of historical road casualties). To control for the RTM effects, the total number of accidents at the site in the before period has been considered in the propensity score model. The propensity score model can match the treated and untreated groups with comparable historical accident records in the pre-treatment period. "Similar" groups can be defined clearly as those with similar propensity scores, thereby avoiding selection bias and RTM and thus ensuring that the difference in the outcomes between the treatment and control groups can be attributed to the treatment. The PSM method also provides a solution to the problem of similarity in the EB method. The EB approach can be improved by using a propensity score matched data sample. Hence, the PSM and improved EB approaches are both employed in this study to ensure validity of the estimation results.

The paper is organized as follows. A literature review is presented in the next section. The methods and data used in this study are described in Sections 3 and 4. The estimation results are presented in Section 5, followed by conclusions in the final section.

2. Literature review

A number of studies have been conducted to examine the safety effects of speed cameras in recent years (Carnis & Blais, 2013; Christie et al., 2003; Gains et al., 2004; Graham et al., 2019; Montella et al., 2015; Mountain et al., 2005; Li et al., 2013; Li & Graham, 2016). For instance, the safety effectiveness of 65 fixed speed cameras on highways in Flanders-Belgium was evaluated (De Pauw et al., 2014). The estimate of the severe crash reduction was 27% (within 250 m from the camera site) and 23% (within 250–500 m from the camera site). In addition, a study based on 38 fixed cameras in Cali, Colombia studied accidents and traffic violations within circles of 250 m of the camera. The reduction in all crashes was estimated to be 5.3% per year (Martínez-Ruíz et al., 2019). Another study of French Automated Speed Enforcement Program (ASEP) concluded that the ASEP was associated with a reduction of 19.7% in traffic fatalities and crashes with injuries (Blais & Carnis, 2015). It has been found that speed cameras can effectively reduce the incidence of crashes. However, there is a lack of studies on how speed camera sites are selected and whether these selection criteria are reasonable.

Regarding the selection of speed cameras sites, most previous studies have focused on identification of crash hotspots, mostly based on crash frequency, rate, and severity (Cheng & Washington, 2005; Manepalli & Bham, 2016; Wang et al., 2018). The objective of identifying hotspots is to find sites (road segments, intersections, ramps, etc.) with potential high risk and these sites can then be considered as candidates for safety countermeasures later. More recently, Wang et al. (2018) applied five hotspot identification methods to find hazardous locations for road segments under municipal jurisdiction in Connecticut. In this study, the roadway hotspots are identified and ranked based on the crash frequency and relative severity index. In addition, the study by Borsos et al. (2016) employed Empirical Bayes method to localize high accident concentration sites in Italy and Hungary. Two conventional indicators, the absolute number of accidents and the accident rate, are adopted for the purpose of site ranking. Although these studies rank the sites based on crash frequency, rate and severity, there is no instruction on how to select appropriate candidates for implementation of safety countermeasures.

Recently, a few studies have been conducted on site selection criteria. For instance, the study performed by Newstead (2016) showed that sites with the highest number of total crashes are viewed as the most appropriate candidates for speed camera implement. Ko et al. (2013) evaluated the impact and site selection criteria for the red light camera systems in Texas in 2008 using an empirical Bayes methodology. This study suggested that the safety effect can vary when the criteria change. There is even a negative effect on road crashes when the cameras sites are selected inappropriately. Another relevant study by Lord and Kuo (2012) examined how site selection criteria influence the estimation of safety effectiveness of treatments. The general idea of site selection bias is that setting entry criteria will convert the original population distribution into a truncated sample distribution leading to changes in the characteristics of the new distribution. Hence, ignoring these changes will create a biased estimator for the safety effectiveness. Their results showed that when estimating the performance of the treatment, higher entry criteria will cause a larger site selection bias, which could lead to larger values of the safety effectiveness.

Regarding methods for road safety evaluation studies, several approaches have been proposed and applied in previous studies (Christie et al., 2003; Guo et al., 2018; Hauer et al., 2002; Wood & Donnell, 2017). The before-after study with control groups is one of the most common approaches to investigate the impacts of speed cameras (Bar-Gera et al., 2017; De Pauw et al., 2014). However, the traditional before-after method usually ignores the RTM. For this reason, EB method has been used as an effective approach in controlling for the limitation of before-after studies (Elvik et al., 2017; Hauer et al., 2002; Høye, 2015). The EB method can be regarded as a statistically defensible means of increasing the precision of estimation and accounting for the RTM bias. However, it is also known that the similarity problem between treated and control units can adversely affect the performance of the EB approach, based on empirical results from previous studies (Lord & Kuo, 2012; Wood & Donnell, 2017). Thus, an alternative approach, the PS method, is proposed to overcome the challenges related to the EB method in the road safety evaluation studies recently (e.g., Hou et al., 2019; Sasidharan & Donnell, 2013;

Wood & Donnell, 2016). The PS method is a causal inference method using the observed covariates of each unit to predict the probabilities that the units received the treatment, which has been widely studied and applied in many evaluations of social, economic and medical programs (e.g., Chattopadhyay et al., 2016; Kowaleski-Jones et al., 2018; Shahidi et al., 2019).

3. Methods

In this section, we first introduce the speed camera sites selection criteria and the scenarios designed in this study. Then the PSM and EB methods are introduced respectively. Both approaches are employed in this study to compare the accuracy and reliability of estimation.

3.1. Selection criteria

The ''Handbook of Rules and Guidance for the National Safety Camera Programme for England and Wales" defines the rules and guidelines for speed camera site selection (DfT, 2005). According to the handbook, the main rules related to the historical crashes data for proposed core sites are:

Criterion 1: number of killed and seriously injured collisions (KSIs): at least 3 KSIs per km in the baseline period. Baseline period is a three-year-period before the installation of all speed cameras, and in our study the baseline period is from 2002 to 2004.

Criterion 2: risk value required: 22 per km. Risk value is introduced in the handbook for new camera sites selection. That is ''new camera sites will be selected using an assessment that includes the level of fatal, serious and slight collisions. The combined level of collisions will be expressed as a numerical scale and assessed relative to the road classification for the site" (DfT, 2005). For instance, fatal or serious injury collision = 5 (i.e. 2 serious collisions = 10), slight injury collision = 1 (i.e. 5 slight collisions = 5).

In this study, the KSI crashes and risk values are calculated over a three-year period before the implementation of speed cameras, and the road sections are categorized into three types according to the treatment status: camera sites meeting and not meeting the selection criteria, and sites without speed cameras. According to the handbook (DfT, 2005), ''a site length is the distance between two points within which collisions, casualties and speeds are measured and camera enforcement takes place." Although, site length is defined in the handbook, there is no description on how the site length is determined when calculating the criteria. In this study, we also evaluated the criteria of site length with different ranges of distance to camera sites on both directions. It is noteworthy that if a site length meets a major junction, the length is terminated. And if a site length overlaps any other site length with an earlier camera deployment date, it will be excluded to avoid double counting or misclassification of before-after status of crashes.

3.2. Scenarios design

This study evaluates the camera site selection criteria by estimating and comparing the effectiveness of speed cameras in various scenarios.

(1) We first compare the characteristics between different types of camera sites. The factors potentially influencing the installment of speed cameras are compared via pair-wise *T*-test.
(2) Then the effectiveness of speed camera sites meeting and not meeting ***criterion 1*** are evaluated, respectively, using the PSM method and EB method. As regards to ***criterion 2***, the same procedure is applied.
(3) Another issue that is not clearly stated in the handbook is which site length should be chosen when calculating ***criterion 1*** and ***2***. It is possible that different length may be used by local authorities, hence affecting the effectiveness of speed cameras. Thus, we calculate the criteria using different site length, including 500 m, 1000 m, and 1500 m. Then the effectiveness of speed camera sites is compared based whether they meet the criteria or not.
(4) Finally, we evaluate and make suggestions on the current criteria by comparing the effects of speed cameras with different possible KSI numbers and risk values.

3.3. Propensity score matching method

The propensity score matching methods are based on the idea that the control or reference group should have similar characteristics with the treated ones. The propensity score can be constructed as a scalar value to account for the probability that a unit is assigned to treated status (i.e., to a speed camera) (Rosenbaum & Rubin, 1983).

The PSM method is applied under the potential outcomes framework. Each unit is associated with two potential outcomes corresponding to two treatment conditions: $D_i = 1$ if unit i receives the treatment (e.g., speed camera sites) and $D_i = 0$ otherwise, where $i = 1, \ldots, N$, and N represents the total number of units. In this study, the treatment groups are further separated into several subgroups according to ***criterion 1*** and ***2***. $Y_i(D)$ represents the potential outcome for unit i. The treatment impact of unit i can be described as:

$$\phi_i = Y_i(1) - Y_i(0) \tag{1}$$

In practice, the parameter of interest is usually the average treatment effect on the treated (ATT), which can be described as:

$$\phi_{ATT} = E(\phi|D=1) = E(Y(1)|D=1) - E(Y(0)|D=1) \tag{2}$$

A fundamental problem is that it is not possible to observe the outcomes of the same unit i for two treatment conditions at the same time (Holland, 1986). Instead, a control group with similar characteristics to the treated units is usually selected to model the counter-factual outcomes.

The propensity score is estimated based on a vector of control covariates X. Conditional on the propensity score, the differences between the treated and untreated units can be accounted for and solely attributed to the treatment effects. Three crucial assumptions need to be satisfied to ensure the validity of the PSM method (Rosenbaum & Rubin, 1983). The first assumption is stable unit treatment value assumption (SUTVA), which defines the treatment assigned to a unit have no impact on the outcomes of others. The second is the conditional independence assumption (CIA), which requires that the potential outcomes are conditionally independent of the treatment assignment conditional on the observable covariates. The last is common support condition (CSC). This assumption is also known as the overlap condition, ensuring the probability of being treated and untreated is positive for the units with the same X values.

It is important to verify these assumptions when assessing the performance of the propensity score matching estimation. The CIA assumes that there should be no statistically significant differences between the covariate means of the treatment and comparison units. Therefore, a balancing test is conducted to check the validity of conditional independence assumption. When verifying the CSC, there are several ways to check the overlap and the region of common support between treatment and control groups and in this study a visual inspection of the propensity score distribution for both groups is presented.

In this paper, a binary logit model is used to estimate the propensity score:

$$P(Y = 1|X) = \frac{\exp(\alpha + \beta X)}{1 + \exp(\alpha + \beta X)} \tag{3}$$

where P is the propensity score for each observation, X is the vector of covariates, α is the intercept, and β is the vector of parameters to be estimated.

Following estimation of the propensity scores, a matching algorithm is selected. Several matching algorithms have been discussed in previous studies, including nearest neighbor matching, kernel and local linear matching, caliper and radius matching, mahalanobis matching and genetic matching. For detailed discussions of these matching algorithms, please refer to the work by Heinrich et al. (2010) and Wood and Donnell (2016).

There is no theoretical rule on how to choose the most appropriate matching algorithm for estimations. Given a large sample, the results from different algorithms should be similar and therefore the choice is not essential. It is good practice to try several matching methods for more credible results. In this study, K-nearest neighbor matching, kernel and local linear matching, and caliper and radius matching are used. A number of programs are available to estimate the treatment effects and psmatch2 in STATA developed by Leuven and Sianesi (2003) program is used in this study.

3.4. Empirical Bayes

EB methods have been widely used in before-after road safety countermeasures evaluation studies (Elvik et al., 2017; Høye, 2015; Wood et al., 2015). A Poisson-Gamma (negative binomial) model is usually applied:

$$z \sim \text{Poisson}(\lambda\varepsilon) \tag{4}$$

$$\ln(\lambda) = a + bX + \varepsilon \tag{5}$$

where z is the observed number of accidents, λ is the expected number of accidents, ε is a Gamma distributed random effect, a, b are the vector of estimable regression parameters and X is the vector of covariates.

Using this model, the predicted number of crashes in the before period, $\hat{M}_B$ can be obtained as

$$\hat{M}_B = \rho\hat{\lambda}_B + (1 - \rho)X_B \tag{6}$$

where X_B is the observed number of crashes in the before period, $\hat{\lambda}_B$ is the EB estimate of total number of crashes based on SPF before treatment, and ρ is a weight factor.

To take account of the trend in accidents between the before and after periods, the predicted accidents in the after period are calculated using a reference group. The estimate of accidents in the after period had treatment not occurred, $\hat{M}_A$, can be evaluated after adjusting the time trend effect using:

$$\hat{M}_A = \left(\frac{N_{A-P}}{N_{B-P}}\right)\hat{M}_B \tag{7}$$

where $N_{A\text{-}P}$ and $N_{B\text{-}P}$ are the numbers of accidents for total population in the before and after periods.

Then the crash number in the after period can be estimated as:

$$\hat{M'}_A = \left(\frac{f_A}{f_{A-P}/f_{B-P} * f_B}\right)^{\beta}\hat{M}_A \tag{8}$$

where f_A is the observed traffic flow, $f_{A\text{-}P}$ and $f_{B\text{-}P}$ are the traffic flow for whole population in the before and after periods, f_B is the observed traffic flow in the before period.

It has been shown that the EB approach can be improved by using a propensity score matched data sample (Li et al., 2013). In this study, the control group is refined via matching. Then the EB method is applied to the refined control group.

4. Data

4.1. Sample size

The available data allow us to observe 332 speed camera sites from 10 English administrative districts, including Cheshire, Dorset, Hertfordshire, Lancashire, Leicester, London, Manchester, Merseyside, Sussex, and Westmiddle. Since the criteria for selecting camera sites in UK have been published in 2004 and improved in 2005 (DfT, 2004, 2005) and the associated safety effects may change as time goes by (Høye, 2015), it is necessary to limit our study to a short period. Therefore, the research period should be appropriate in order to control the impact resulted from the time span. In this study, only speed cameras installed between 2005 and 2007 are chosen due to the time varying effects and the criteria published time. In addition, as suggested in previous studies (Gains et al., 2004, 2005; Høye, 2015; Li & Graham, 2016; Mountain et al., 2005), accident data from 2002 to 2010 are used to ensure that three years of data are observed before and after the installment of all speed camera sites.

It is worth noting that a sufficient number of control candidates should be included to guarantee matching quality (Kurth et al., 2006; Peikes et al., 2008). A total number of 2,513 potential control sites are included in this study, which are chosen randomly within the 10 districts mentioned before. These sites have no camera before 2010 to avoid misclassification of before-after status of crashes and are at least 2 km away from any camera sites to avoid double counting of crashes and the treatment effects. Fig. 1 shows an example of treatment and control sites (site length = 500 m) in London.

4.2. Covariates

Both factors affecting the treatment assignment and safety effects of speed cameras are included in the models (Gains et al., 2005; Høye, 2015; Li & Graham, 2016; Mountain et al., 2005):

(1) KSIs: the number of killed and seriously injured collisions in the baseline years (total number between 2002 and 2004)
(2) PICs: the number of personal injury collisions in the baseline years (total number between 2002 and 2004)
(3) AADF: the annual average daily traffic flow around the site
(4) Speed limits: focusing on sites with speed limits of 30 mph and 40 mph throughout the UK
(5) Junctions/km: the number of minor road junctions per kilometer within the site length
(6) Road class: e.g. A road, B road, Minor road

In addition, site selection should be centered on locations where speed is a contributing factor in crashes. And it is necessary to assess sites where speeding is particularly hazardous to road users (NCHRP Report 729, 2012). Thus, we further include covariates such as population density, school zones, percentage of residential area, and the index of multiple deprivation (IMD) to compare the factors of different types of sites. The IMD integrates data on the following seven deprivation domain indices into one overall deprivation score: income deprivation, employment deprivation, education, skills and training deprivation, health deprivation and disability, crime, barriers to housing and services, and living environment deprivation. IMD has previously been shown to be

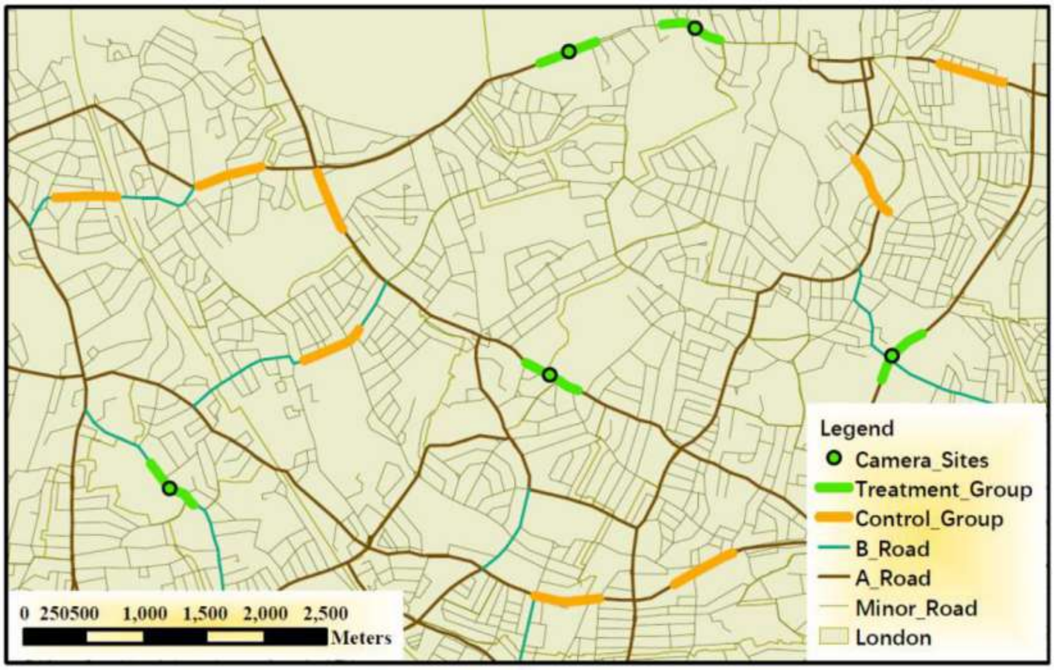


Fig. 1. An example of treatment and control sites (site length = 500 m) in London.

positively related to road traffic casualties (e.g., Dissanayake et al., 2009; Feleke et al., 2018; Graham & Stephens, 2008). In summary, the covariates included are shown in Table 1.

5. Results

In this section, the following issues are discussed: (1) the comparison of the characteristics of speed camera sites meeting and not meeting the criteria; (2) the difference in the safety effectiveness of speed camera sites meeting and not meeting the criteria; (3) the optimal site length when calculating the criteria; and (4) the optimum threshold values for ***criterion 1*** and ***2*** to achieve maximum safety effects of speed cameras.

5.1. Comparison of the sites characteristics

The characteristics of different types of sites are compared via a pair-wise *T*-test. Table 2 shows the results for speed camera sites meeting and not meeting the criteria. Significant differences can be observed for all covariates except speed limits for both ***criterion 1*** and ***2***. This is probably because the potential sites for speed camera installment are usually those with a speed limit of 30 mph and 40 mph throughout the UK. And about 81.1% of speed cameras are installed on roads with a speed limit of 30 mph. As shown in Table 2, there are critical differences in the percentage of A road, the number of junctions, the percentage of sites in urban area, the percentage of residential area, and population density. The camera sites meeting the criteria tend to be located on roads with more complex environment and densely populated areas. In addition, proximity to schools is a major consideration for both sites meeting and not meeting the criteria.

Table 3 presents the differences between the sites not meeting the selection criteria and the control sites. According to the handbook (DfT, 2005), in addition to sites that meet the criteria, speed cameras can be also installed on roads "where the local community requests the partnership enforce at a particular site because traffic speeds there are causing concern for road safety." As shown in Table 3, there are significant differences in factors including IMD, percentage of residential area, population density, and school zones, all of which are related to the community concern.

Table 1
Covariates included in the significance test.

Covariates	Description
A road (%)	Percentage of A roads
B road (%)	Percentage of B roads
Speed limit 30 mph (%)	Percentage of sites with the speed limit of 30 mph
Speed limit 40 mph (%)	Percentage of sites with the speed limit of 30 mph
AADF	The annual average daily traffic around the site
Junction	The number of road minor junctions per kilometer within a effective length of the site
Urban (%)	Percentage of sites in urban area
Rural (%)	Percentage of sites in rural area
IMD	The index of multiple deprivation
Residential (%)	Percentage of residential area
Population density	Residential population per km^2
School zones (%)	Percentage of sites within school zones

5.2. Estimation of propensity scores

The results in Table 4 show that the covariates included in the propensity score model are all significant. As Westreich et al. (2011) suggested, however, the main purpose of the PSM is not to predict the treatment assignment, but to control for confounding via balancing the control covariates. So, it is important to test the matching quality.

Two approaches are used to examine the validity of the PSM method, one of which is a visual inspection of the propensity score distributions for both the treatment and control groups. The histograms in Fig. 2 presents an example of the distributions of the propensity scores, indicating the overlap assumption is plausible. There are 332 speed camera sites and 2,513 potential control sites. Depending on different scenarios, the ratio of the number of

Table 2
Comparison of the sites characteristics (sites meeting criteria VS sites not meeting criteria).

	Criterion 1 (KSIs ≥ 3)					Criterion 2 (Risk Value ≥ 22)				
	Not meet	meet	Diff	Std. Err.	*T*-test	Not meet	meet	diff	Std. Err.	*T*-test
A road (%)	61.33	87.35	−26.02*	5.41	−4.81	66.67	82.81	−16.15*	6.31	−2.56
B road (%)	23.68	5.99	17.70*	4.29	4.13	19.61	9.38	10.23*	5.01	2.04
Speed limit 30mph (%)	84.13	83.23	0.90	5.54	0.16	77.50	84.78	−7.28	6.49	−1.12
Speed limit 40mph (%)	14.29	14.29	0.00	5.22	0.00	22.50	12.50	10.00	6.10	1.64
AADF	28,267	30,628	−2361	3126	−0.76	22,886	31,750	−8864*	3517	−2.52
Junction	6.39	10.17	−3.77*	0.61	−6.14	5.43	9.93	−4.50*	0.69	−6.49
Urban (%)	82.89	96.41	−13.51*	3.63	−3.73	78.43	95.83	−17.40*	4.10	−4.25
Rural (%)	17.11	3.59	13.51*	3.63	3.73	21.57	4.17	17.40*	4.10	4.25
IMD	23.70	29.20	−5.51*	2.30	−2.40	18.16	29.96	−11.80*	2.54	−4.65
Residential (%)	57.89	70.06	−12.17**	6.52	−1.87	49.02	70.83	21.81*	7.35	−2.97
Population density	2813.3	6848.0	−4034.7*	547.6	−7.37	2033.6	6529.8	−4496.1*	626.4	−7.18
School zones (%)	64.47	89.22	−24.75*	5.16	−4.80	54.90	88.54	−33.64*	5.75	−5.85

Notes: Figures are significant at: *95%, **90%.

Table 3
Comparison of the sites characteristics (sites not meeting criteria VS untreated sites).

	Criterion 1 (KSIs ≥ 3)					Criterion 2 (Risk Value ≥ 22)				
	Not meet	Untreated	Diff	Std. Err.	*T*-test	Not meet	Untreated	diff	Std. Err.	*T*-test
A road (%)	61.84	99.08	−37.24*	1.98	−18.81	66.67	99.07	−32.41*	2.09	−15.47
B road (%)	23.68	0.92	22.76*	1.81	12.56	19.61	0.93	18.68*	1.91	9.80
Speed limit 30mph (%)	77.63	51.61	26.02*	5.91	4.40	68.63	51.45	17.18*	7.18	2.39
Speed limit 40mph (%)	13.16	19.20	−6.04	4.67	−1.29	19.61	19.24	0.37	5.69	0.07
AADF	28,267	20,983	7284*	2133	3.41	22,886	20,973	1913	2455	0.78
Junction	6.39	7.07	−0.68	0.69	−0.98	5.43	7.10	−1.67*	0.84	−1.99
Urban (%)	85.14	70.68	14.45*	5.43	2.66	78.43	70.68	7.75	6.53	1.19
Rural (%)	14.86	29.32	−14.45*	5.43	−2.66	21.57	29.32	−7.75	6.53	−1.19
IMD	24.34	20.98	3.36**	1.96	1.72	18.16	20.98	−2.82	2.30	−1.22
Residential (%)	59.46	34.30	25.16*	5.77	4.36	49.02	34.30	14.72*	6.87	2.14
Population density	2813.3	2008.7	804.6*	276.2	2.91	2033.6	2014.1	19.5	329.8	0.06
School zones (%)	64.47	28.51	35.97*	5.43	6.62	54.90	28.62	26.28*	6.56	4.01

Notes: Figures are significant at: *95%, **90%.

Table 4
Results of the propensity score model.

Treatment	Coef.	Std. Err.	z	P > z	[95% Conf.	Interval]
PICs in baseline years	0.069	0.019	3.69	<0.001	0.032	0.106
KSIs in baseline years	−0.787	0.151	−5.21	<0.001	−1.084	−0.491
AADT in baseline years	3.30E-05	7.54E-06	4.38	<0.001	1.82E-05	4.77E-05
Junction	0.041	0.025	3.39	0.001	−0.008	0.089
A road	2.777	0.362	−7.67	<0.001	−3.487	−2.067
B road	−1.373	0.399	−3.44	0.001	−2.154	−0.591
Speed limit 30mph	1.070	0.418	2.56	0.010	0.251	1.889
Speed limit 40mph	1.383	0.452	3.06	0.002	0.497	2.269
_cons	−1.865	0.501	−3.72	<0.001	−2.847	−0.882

control candidates to the treated ones ranges from 6:1 to over 10:1 in order to obtain a sufficient overlapping.

It is also recommended that a balance test be conducted to check the validity of conditional independence assumption. Theoretically, there should be no significant differences in the covariate means between the treated and untreated groups after matching. Table 5 shows an example of the *t*-test of the differences in covariate means before and after matching. The results show that significant differences can be observed for all covariates before matching and all covariates are well balanced after matching.

5.3. Evaluating speed camera sites selection criteria

In this section, we investigate how speed camera sites are selected and evaluate whether these selection criteria are reasonable and effective. The safety effects of three types of sites are evaluated and compared by the PSM and EB methods. In addition, different site length is used to calculate the criteria. Finally, alternative threshold values for ***criterion 1*** and ***2*** are compared in order to find optimal criteria.

Table 6 and Table 7 show the estimation results of the effects of speed cameras on annual PICs and KSIs per kilometer for ***criterion 1*** and ***2,*** respectively. The estimation results by the PSM method based on different matching algorithms and EB method are very similar. In Table 6, the reduction in the absolute number of PICs ranges from 0.597 to 1.147 per km per year, while the number varies from 0.313 to 0.357 for sites not meeting the criterion. Similarly, in Table 7, the average reductions are 0.845 for up to 500 meters, 0.767 for up to 1,000 meters and 0.569 for up to 1,500 meters respectively in the annual PICs per km. For the sites not meeting ***criterion 2***, the reductions in the annual PICs per km are 0.293 for up to 500 meters, 0.203 for up to 1,000 meters and 0.186 for up to 1,500 meters, respectively.

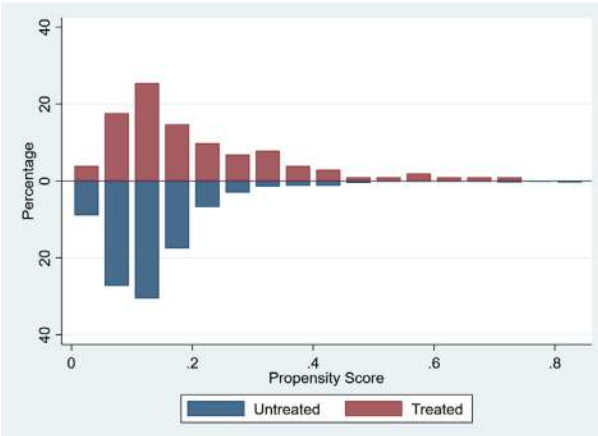


Fig. 2. Propensity score distribution.

Table 8 shows the average effects of speed cameras on annual PICs per km in percentages. Similarly, the results suggest that the speed camera sites are most effective when the criteria are calculated using a road length of 500 meters (15.63% for ***criterion 1*** and 16.85% for ***criterion 2***). For both ***criterion 1*** and ***2***, the speed camera sites meeting the criteria perform better in reducing the annual PICs per km in percentages. The results suggest that the speed cameras are more effective when the selection criteria are met. In addition, a road length of 500 meters is recommended when calculating ***criterion 1*** and ***2***.

Table 9 and Fig. 3 present the safety effects of speed cameras at sites with different number of KSIs in the baseline years. For the reduction in annual PICs per km, the estimates ranges from 0.653 for sites with more than one KSI to 0.838 for sites with more than six KSIs. In terms of the effects on KSIs, the annual reduction per km in KSIs ranges from 0.232 to 0.297. In terms of the reduction in percentages, however, the results suggest that the speed cameras are most effective at sites with at least 3 KSIs in the baseline years, where the reduction is 13.20%.

Table 10 and Fig. 4 show the average effects of speed cameras given different risk values. The results also show increasing effects on reducing the absolute number of PICs and KSIs when the risk values increase. The annual reductions in PICs vary from 0.614 to 1.149 per km for different risk values, while the annual reductions per km in KSIs range from 0.220 to 0.285. In terms of the reduction in percentages, however, the speed cameras are most effective at

Table 5
Checking the covariates balance between groups before and after using nearest neighbors ($k = 3$) matching.

Covariates	Sample	Mean		%bias	%reduced	*t*-test	
		Treated	Control		|bias|	*t*	*p* > |*t*|
PICs in baseline years	Unmatched	48.866	25.66	90.1		10.29	<0.001
	Matched	48.866	50.955	−8.1	91	−0.53	0.594
KSIs in baseline years	Unmatched	8	3.6643	120.3		13.57	<0.001
	Matched	8	7.9851	0.4	99.7	0.03	0.977
AADT in baseline years	Unmatched	30,239	24,050	38		4.21	<0.001
	Matched	30,239	30,499	−1.6	95.8	−0.11	0.916
Junction	Unmatched	10.791	8.1804	49.7		5.01	<0.001
	Matched	10.791	10.575	4.1	91.7	0.35	0.73
A road	Unmatched	0.7228	0.9487	−63.8		−8.44	<0.001
	Matched	0.7228	0.7525	−8.4	86.9	−0.48	0.633
B road	Unmatched	0.1881	0.0443	45.9		5.88	<0.001
	Matched	0.1881	0.1485	12.6	72.5	0.75	0.454
Speed limit 30mph	Unmatched	0.8438	0.7122	32		3.77	<0.001
	Matched	0.8438	0.8646	−5.1	84.2	−0.58	0.564
Speed limit 40mph	Unmatched	0.1198	0.1775	−16.3		−1.95	0.052
	Matched	0.1198	0.1302	−2.9	82	−0.31	0.758

Table 6
Average effects of speed cameras on annual PICs/KSIs per km in absolute number for criterion 1.

	Criterion 1 (KSIs $\geq$ 3)					
	Meet criterion 1			Not meet criterion 1		
	500 m	1000 m	1500 m	500 m	1000 m	1500 m
Changes in annual PICsr per km in absolute number						
Unmatched	−2.882*	−2.049*	−1.652*	−0.368	−0.633**	−0.368
K-nearest Neighbors Matching ($K = 1$)	−1.357*	−0.871*	−0.611**	−0.438	−0.397	−0.338
K-nearest Neighbors Matching ($K = 3$)	−1.181*	−0.856*	−0.589**	−0.394	−0.231	−0.294
Kernel Matching (Bandwidth = 0.05)	−1.011*	−0.807**	−0.593**	−0.303	−0.314	−0.303
Radius Matching (Caliper = 0.05)	−1.089*	−0.819*	−0.605**	−0.321	−0.332	−0.321
EB	−1.093*	−0.828*	−0.59	−0.327	−0.315	−0.311
Average Effect	−1.147	−0.836	−0.597	−0.357	−0.318	−0.313
Changes in annual KSIs per km in absolute number						
Unmatched	−1.742*	−0.833*	−0.508*	0.147	−0.141	0.035
K-nearest Neighbors Matching ($K = 1$)	−0.622*	−0.279*	−0.239*	0.067	−0.108	0.034
K-nearest Neighbors Matching ($K = 3$)	−0.635*	−0.235*	−0.241*	0.158	−0.100	0.025
Kernel Matching (Bandwidth = 0.05)	−0.637*	−0.270*	−0.214**	0.141	−0.114	0.021
Radius Matching (Caliper = 0.05)	−0.687*	−0.278*	−0.223*	0.147	−0.122	0.027
EB	−0.612*	−0.256*	−0.205*	0.124	−0.114	0.023
Average Effect	−0.638	−0.264	−0.224	0.128	−0.112	0.026

Notes: Figures are significant at: *95%, **90%.

Table 7
Average effects of speed cameras on annual PICs/KSIs per km in absolute number for criterion 2.

	Criterion 2 (Risk value ≥ 22)					
	Meet criterion 2			Not meet criterion 2		
	500 m	1000 m	1500 m	500 m	1000 m	1500 m
Changes in annual PICsr per km in absolute number						
Unmatched	−2.278*	−2.082*	−1.886*	−0.238	−0.591**	0.102
K-nearest Neighbors Matching (*K* = 1)	−0.951*	−0.740*	−0.687**	−0.209	−0.192	−0.260
K-nearest Neighbors Matching (*K* = 3)	−0.815*	−0.734**	−0.549**	−0.344	−0.166	−0.119
Kernel Matching (Bandwidth = 0.05)	−0.814*	−0.803*	−0.548**	−0.320	−0.221	−0.194
Radius Matching (Caliper = 0.05)	−0.830*	−0.811 *	−0.513**	−0.314	−0.229**	−0.185
EB	−0.812*	−0.747 *	−0.547	−0.275	−0.208	−0.168
Average Effect	−0.845	−0.767	−0.569	−0.293	−0.203	−0.186
Changes in annual KSIs per km in absolute number						
Unmatched	−1.165*	−0.707*	−0.477*	0.020	−0.107	−0.047
K-nearest Neighbors Matching (*K* = 1)	−0.402*	−0.224**	−0.193**	0.020	−0.009	−0.027
K-nearest Neighbors Matching (*K* = 3)	−0.539*	−0.213*	−0.183**	−0.024	−0.024	−0.018
Kernel Matching (Bandwidth = 0.05)	−0.528*	−0.229*	−0.156**	0.012	−0.028	−0.041
Radius Matching (Caliper = 0.05)	−0.550*	−0.233*	−0.157**	0.024	−0.029	−0.034
EB	−0.491*	−0.228*	−0.165**	−0.081	−0.027	−0.027
Average Effect	−0.502	−0.226	−0.171	−0.010	−0.024	−0.029

Notes: Figures are significant at: *95%, **90%.

Table 8
Average effects of speed cameras on annual PICs per km in percentage.

	Changes in annual PICs per km in Percentage					
	500 m	1000 m	1500 m	500 m	1000 m	1500 m
	Criterion 1 (KSIs ≥ 3)			KSIs < 3		
Unmatched	−20.99*	−19.46*	−17.72*	−18.32*	−18.44*	−15.32*
K-nearest Neighbors Matching (*K* = 1)	−16.10*	−13.70*	−7.84**	−12.42*	−9.47**	−8.29**
K-nearest Neighbors Matching (*K* = 3)	−17.62*	−12.22*	−7.38**	−11.45*	−8.40**	−8.51**
Kernel Matching (Bandwidth = 0.05)	−14.31*	−13.22*	−8.64 **	−12.72*	−10.86**	−8.51**
Radius Matching (Caliper = 0.05)	−14.50*	−13.65*	−8.80*	−12.87 *	−11.53*	−8.70*
Average	−15.63	−13.20	−8.17	−12.37*	−10.06	−8.51
	Criterion 2 (Risk value ≥ 22)			Risk value < 22		
Unmatched	−21.12*	−19.15*	−18.28*	−17.89*	−17.97*	−17.63*
K-nearest Neighbors Matching (*K* = 1)	−16.67*	−12.57*	−9.81**	−10.24	−10.61	−10.16*
K-nearest Neighbors Matching (*K* = 3)	−17.69*	−11.72*	−9.60	−11.68	−10.23	−9.38**
Kernel Matching (Bandwidth = 0.05)	−16.20*	−12.52*	−9.74	−12.01	−10.20	−10.95*
Radius Matching (Caliper = 0.05)	−16.86*	−11.94*	−10.22**	−12.81**	−11.14	−10.07*
Average	−16.85	−12.19	−9.84	−11.68	−10.54	−10.14

Notes: Figures are significant at: *95%, **90%.

Table 9
Average effects of speed cameras by KSIs in the baseline years.

The number of KSIs	≥1	≥2	≥3	≥4	≥5	≥6
Changes in annual PICs per km in absolute number						
Unmatched	−1.447*	−1.777*	−2.049*	−2.390*	−2.550*	−2.731*
K-nearest Neighbors Matching (*K* = 1)	−0.528**	−0.834*	−0.871*	−0.907*	−0.818**	−0.842**
K-nearest Neighbors Matching (*K* = 3)	−0.790*	−0.547**	−0.856*	−0.857*	−0.827**	−0.827**
Kernel Matching (Bandwidth = 0.05)	−0.637*	−0.851*	−0.807*	−0.795*	−0.851**	−0.846**
Radius Matching (Caliper = 0.05)	−0.656*	−0.835*	−0.819*	−0.797*	−0.865*	−0.835**
Average Effect	−0.653	−0.767	−0.838	−0.839	−0.840	−0.838
Changes in annual KSIs per km in absolute number						
Unmatched	−0.615*	−0.715*	−0.833*	−0.942*	−1.037*	−1.099*
K-nearest Neighbors Matching (*K* = 1)	−0.233*	−0.182**	−0.279*	−0.321*	−0.303*	−0.294**
K-nearest Neighbors Matching (*K* = 3)	−0.231*	−0.223*	−0.235*	−0.262**	−0.287**	−0.293**
Kernel Matching (Bandwidth = 0.05)	−0.233*	−0.271*	−0.270*	−0.288**	−0.291**	−0.296**
Radius Matching (Caliper = 0.05)	−0.233*	−0.272*	−0.278*	−0.299*	−0.294*	−0.303*
Average Effect	−0.232	−0.237	−0.266*	−0.293	−0.294	−0.297
Changes in annual PICs per km in percentage						
Unmatched	−17.68*	−19.65*	−19.46*	−18.84*	−16.82*	−14.32*
K-nearest Neighbors Matching (*K* = 1)	−11.56*	−11.03*	−13.70*	−9.96**	−7.58*	−5.87
K-nearest Neighbors Matching (*K* = 3)	−12.09*	−10.31*	−12.22*	−9.95*	−7.21*	−5.94**
Kernel Matching (Bandwidth = 0.05)	−11.50*	−13.21*	−13.22*	−10.67*	−7.14**	−5.53
Radius Matching (Caliper = 0.05)	−11.77*	−13.74*	−13.65*	−11.38*	−7.71**	−5.75**
Average Effect	−11.73	−12.07	−13.20	−10.49	−7.41	−5.77

Notes: Figures are significant at: *95%, **90%.

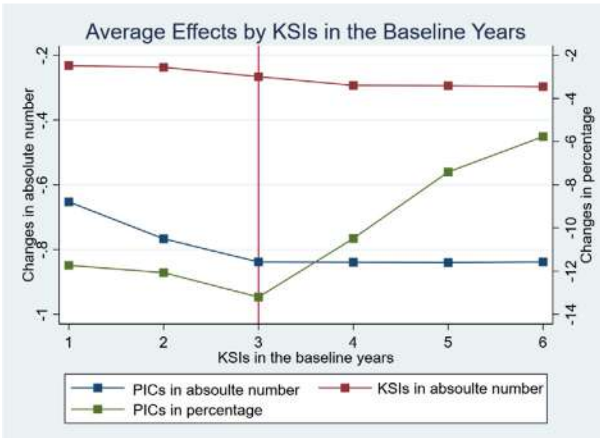


Fig. 3. Average effects by KSIs in the baseline years.

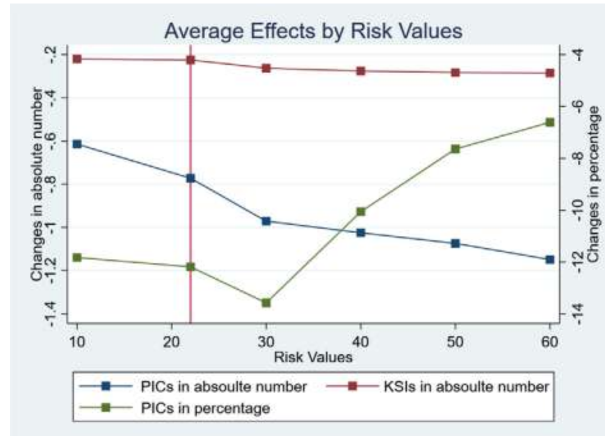


Fig. 4. Average effects by risk values.

sites with a risk value greater than or equal to 30, where the estimate is approximately 13.58%.

6. Discussion and conclusion

Speed camera is an important road safety intervention that has been found to be effective in regulating driving speeds and reducing casualties. However, there is limited research on issues surrounding the selection criteria for speed camera sites, including how these selection criteria are decided and whether the current criteria can achieve optimum safety benefits. This study evaluate the current speed camera sites selection criteria in the UK under different scenarios via propensity score matching and empirical Bayes methods.

In terms of the characteristics between different types of sites, the pairwise comparison results indicate that factors including IMD, percentage of residential area, population density, and school zones are the main considerations when selecting speed camera sites. It is suggested that speed camera sites are more likely to be those with more complex environment and in densely populated areas, which are consistent with the rules about selecting exceptional sites with community concern in the handbook (DfT, 2005). In addition, a comparison of the safety effects of speed camera sites meeting and not meeting the criteria is conducted. The results show that speed camera sites are more effective in reducing crashes when the selection criteria are satisfied and a road length of 500 m should be used by the local authorities to achieve the optimum safety effects of speed cameras.

We further investigate the possible alternative critical values for ***criterion 1*** and ***2***. The results show that as the number of KSIs and the risk value in the baseline years increases, the effects of speed cameras in reducing the absolute number of road casualties also increase. In terms of the reduction in percentages, however, the speed cameras achieve the optimal effect at sites with at least 3 KSIs in the baseline period. This finding is consistent with ***criterion 1***. However, for ***criterion 2***, the speed cameras are most effective when the risk value is greater than or equal to 30, which is higher than the current value of 22.

In summary, this paper contributes to the literature by evaluating the existing speed camera sites selection criteria in the UK. Several critical issues of “Handbook of Rules and Guidance for the National Safety Camera Programme for England and Wales” have

Table 10
Average effects of speed cameras by risk value.

Risk value	≥10	≥22	≥30	≥40	≥50	≥60
Changes in annual PICs per km in absolute number						
Unmatched	−1.697*	−2.082*	−2.347*	−2.692*	−2.879*	−3.117*
K-nearest Neighbors Matching (*K* = 1)	−0.502**	−0.740**	−0.949*	−0.999*	−0.983*	−1.152*
K-nearest Neighbors Matching (*K* = 3)	−0.629*	−0.734**	−0.954*	−0.922*	−1.114*	−1.095*
Kernel Matching (Bandwidth = 0.05)	−0.656*	−0.803*	−0.976*	−1.094*	−1.115*	−1.157*
Radius Matching (Caliper = 0.05)	−0.668*	−0.811*	−1.003*	−1.087*	−1.085*	−1.194*
Average Effect	−0.614	−0.772	−0.971	−1.025	−1.074	−1.149
Changes in annual KSIs per km in absolute number						
Unmatched	−0.621*	−0.707*	−0.762*	−0.883*	−0.922*	−0.957*
K-nearest Neighbors Matching (*K* = 1)	−0.226*	−0.224**	−0.237*	−0.271*	−0.274*	−0.286**
K-nearest Neighbors Matching (*K* = 3)	−0.222*	−0.213*	−0.254*	−0.277*	−0.288*	−0.281*
Kernel Matching (Bandwidth = 0.05)	−0.213*	−0.229*	−0.279*	−0.280*	−0.281*	−0.286*
Radius Matching (Caliper = 0.05)	−0.218*	−0.233*	−0.281*	−0.276*	−0.287**	−0.288
Average Effect	−0.220	−0.225	−0.263	−0.276	−0.283	−0.285
Changes in annual PICs per km in percentage						
Unmatched	−19.48*	−19.15*	−20.30*	−18.80*	−16.47*	−14.79*
K-nearest Neighbors Matching (*K* = 1)	−11.14*	−12.57*	−13.45*	−9.89*	−8.05**	−7.15*
K-nearest Neighbors Matching (*K* = 3)	−11.97*	−11.72*	−13.69*	−9.60*	−7.58**	−6.19**
Kernel Matching (Bandwidth = 0.05)	−11.86*	−12.52*	−13.48*	−10.63*	−7.44	−6.46
Radius Matching (Caliper = 0.05)	−12.34*	−11.94*	−13.72*	−10.10*	−7.48	−6.66**
Average Effect	−11.83	−12.19	−13.58	−10.06	−7.64	−6.61

Notes: Figures are significant at: *95%, **90%.

been discussed in the paper, which will be helpful to policy-making and legislation by providing new evidences and practical recommendations on enhancing the effectiveness of speed camera assignment. First, this study identifies the characteristics of speed camera sites not meeting the criteria, which can help to understand the community concern when selecting potential camera sites. Second, an optimal site length is suggested for the local authorities when calculating the criteria. Third, we propose an optimum threshold value for risk value to achieve maximum safety effects when selecting the speed camera sites.

There are also some limitations in this study. First, only the fixed speed camera sites selection criteria in the UK are analyzed. Due to the data restriction in this study, it is difficult to evaluate the criteria for selecting mobile speed camera sites. Second, only 30 mph and 40 mph speed limitation are taken into considerations. In fact, the results may be different on the road with a higher speed limitation because cameras at locations with a lower speed limit generate greater effects (De Pauw et al., 2014). In addition, many studies have been limited to the modelling of safety benefits. This study may contribute to the improvements in economic performance of speed cameras generally. Further research could focus on the evaluation of the cost-effectiveness of current speed camera deployment strategies, which is difficult to conduct in this study due to the data restriction.

Acknowledgments

This work was supported by the National Key R&D Program of China (No. 2018YFE0102700).

References

Bar-Gera, H., Schechtman, E., & Musicant, O. (2017). Evaluating the effect of enforcement on speed distributions using probe vehicle data. *Transportation Research Part F, 46*, 271–283.

Blais, E., & Carnis, L. (2015). Improving the safety effect of speed camera programs through innovations: Evidence from the French experience. *Journal of Safety Research, 55*, 135–145.

Borsos, A., Cafiso, S., Agostino, C. D., & Miletics, D. (2016). Comparison of Italian and Hungarian black spot ranking. *Transportation Research Procedia, 14*, 2148–2157.

Carnis, L., & Blais, E. (2013). An assessment of the safety effects of the French speed camera program. *Accident Analysis & Prevention, 51*, 301–309.

Chattopadhyay, A. S., Lin, Y., Hsieh, A., Chang, C., Lian, L., & Fann, C. (2016). Using propensity score adjustment method in genetic association studies. *Computational Biology and Chemistry, 62*, 1–11.

Cheng, W., & Washington, S. P. (2005). Experimental evaluation of hotspot identification methods. *Accident Analysis & Prevention, 37*(5), 870–881.

Christie, S. M., Lyons, R. A., Dunstan, F. D., & Jones, S. J. (2003). Are mobile speed cameras effective? A controlled before and after study. *Injury Prevention, 9*(4), 302–306.

De Pauw, E., Daniels, S., Brijs, T., Hermans, E., & Wets, G. (2014). An evaluation of the traffic safety effect of fixed speed cameras. *Safety Science, 62*, 168–174.

Department for Transport (2004). Handbook of Rules and Guidance for the National Safety Camera Programme for England and Wales for 2005/06.

Department for Transport (2005). Handbook of Rules and Guidance for the National Safety Camera Programme for England and Wales for 2006/07.

Dissanayake, D., Aryaija, J., & Wedagama, D. M. P. (2009). Modelling the effects of land use and temporal factors on child pedestrian casualties. *Accident Analysis and Prevention, 41*(5), 1016–1024.

Elvik, R., Ulstein, H., Wifstad, K., Syrstad, R. S., Seeberg, A. R., Gulbrandsen, M. U., & Welde, M. (2017). An Empirical Bayes before-after evaluation of road safety effects of a new motorway in Norway. *Accident Analysis and Prevention, 108*, 285–296.

Feleke, R., Scholes, S., Wardlaw, M., & Mindell, J. S. (2018). Comparative fatality risk for different travel modes by age, sex, and deprivation. *Journal of Transport & Health, 8*, 307–320.

Gains, A., Heydecker, B., Shrewsbury, J., & Robertson, S. (2004). *The National Safety Camera Programme 3-year evaluation report*. London: PA Consulting Group and UCL for Department for Transport.

Gains, A., Heydecker, B., Shrewsbury, J., & Robertson, S. (2005). *The National Safety Camera Programme 4-year evaluation report*. London: PA Consulting Group and UCL for Department for Transport.

Graham, D. J., & Stephens, D. A. (2008). Decomposing the impact of deprivation on child pedestrian casualties in England. *Accident Analysis & Prevention, 40*(4), 1351–1364.

Graham, D. J., Naik, C., McCoy, E. J., & Li, H. (2019). Do speed cameras reduce road traffic collisions? *PLoS ONE, 14*(9) e0221267.

Guo, Y., Liu, P., Wu, Y., & Chen, J. (2018). Evaluating how right-turn treatments affect right-turn-on-red conflicts at signalized intersections. *Journal of Transportation Safety and Security, 12*, 419–440. https://doi.org/10.1080/19439962.2018.1490368.

Hauer, E., Harwood, D. W., Council, F. M., & Griffth, M. S. (2002). The empirical Bayes method for estimating safety: A tutorial. *Transportation Research Record: Journal of the Transportation Research Board, 1784*, 126–131.

Heinrich, C., Maffioli, A., & Vázquez, G. (2010). A primer for applying propensity-score matching. *Inter-American Development Bank, Technical Notes No. IDB-TN-161*.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81*, 945–960.

Hou, Q., Meng, X., Huo, X., Cheng, Y., & Leng, J. (2019). Effects of freeway climbing lane on crash frequency: Application of propensity scores and potential outcomes. *Physica A, 517*, 246–256.

Høye, A. (2015). Safety effects of section control—An empirical Bayes evaluation. *Accident Analysis and Prevention, 82*, 263–269.

Kowaleski-Jones, L., Zick, C., Smith, K. R., Brown, B., Hanson, H., & Fan, J. (2018). Walkable neighborhoods and obesity: Evaluating effects with a propensity score approach. *SSM - Population Health, 6*, 9–15.

Ko, M., Geedipally, S. R., & Walden, T. D. (2013). Effectiveness and site selection criteria for red light camera systems. *Transportation Research Record: Journal of the Transportation Research Board, 2327*, 53–60.

Kurth, T., Walker, A. M., Glynn, R. J., Chan, K. A., Gaziano, J. M., Berger, K., & Robins, J. M. (2006). Results of multivariable logistic regression, propensity matching propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. *American Journal of Epidemiology, 163*, 262–270.

Leuven, E., & Sianesi, B. (2003). *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing, statistical software components S432001*. Boston College Department of Economics.

Li, H., & Graham, D. J. (2016). Heterogeneous treatment effects of speed cameras on road safety. *Accident Analysis and Prevention, 97*, 153–161.

Li, H., Graham, D. J., & Majumdar, A. (2013). The impacts of speed cameras on road accidents: An application of propensity score matching methods. *Accident Analysis and Prevention, 60*, 148–157.

Lord, D., & Kuo, P. (2012). Examining the effects of site selection criteria for evaluating the effectiveness of traffic safety countermeasures. *Accident Analysis and Prevention, 47*, 52–63.

Manepalli, U. R. R., & Bham, G. H. (2016). An evaluation of performance measures for hotspot identification. *Journal of Transportation Safety and Security, 8*(4), 327–345.

Montella, A. (2010). A comparative analysis of hotspot identification methods. *Accident Analysis and Prevention, 42*, 571–581.

Martínez-Ruíz, D. M., Fandiño-Losada, A., Ponce de Leon, A., Arango-Londoño, D., Mateus, J. C., Jaramillo-Molina, C., Bonilla-Escobar, F. J., Vivas, H., Vanlaar, W., & Gutiérrez-Martínez, M. I. (2019). Impact evaluation of camera enforcement for traffic violations in Cali, Colombia, 2008–2014. *Accident Analysis and Prevention, 125*, 267–274.

Montella, A., Imbriani, L. L., Marzano, V., & Mauriello, F. (2015). Effects on speed and safety of point-to-point speed enforcement systems: Evaluation on the urban motorway A56 Tangenziale di Napoli. *Accident Analysis and Prevention, 75*, 164–178.

Mountain, L. J., Hirst, W. M., & Maher, M. J. (2005). Are speed enforcement cameras more effective than other speed management measures? *Accident Analysis and Prevention, 37*, 742–754.

Newstead, S. V., (2016). Outcomes of a workshop to determine criteria for placement of fixed speed cameras and red light speed cameras under the automated traffic enforcement program in Western Australia. Retrieved from https://www.rsc.wa.gov.au/RSC/media/Documents/Resources/Publications/criteria-for-placement-of-fixed-speed-cameras.pdf.

Oviedo-Trespalacios, O., Truelove, V., Watson, B., & Hinton, J. A. (2019). The impact of road advertising signs on driver behaviour and implications for road safety: A critical systematic review. *Transportation Research Part A: Policy and Practice, 122*, 85–98.

Peikes, D. N., Moreno, L., & Orzol, S. M. (2008). Propensity score matching: a note of caution for evaluators of social programs. *The American Statistician, 62*, 222–231.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41–55.

Sasidharan, L., & Donnell, E. T. (2013). Application of propensity scores and potential outcomes to estimate effectiveness of traffic safety countermeasures: Exploratory analysis using intersection lighting data. *Accident Analysis and Prevention, 50*, 539–553.

Shahidi, F. V., Muntaner, C., Shankardass, K., Quiñonez, C., & Siddiqi, A. (2019). The effect of unemployment benefits on health: A propensity score analysis. *Social Science & Medicine, 226*, 198–206.

Transportation Research Board (2012). NCHRP Report 729: Automated Enforcement for Speeding and Red Light Running.

Wang, K., Zhao, S., Ivan, J. N., Ahmed, I., & Jackson, E. (2018). Evaluation of hot spot identification methods for municipal roads. *Journal of Transportation Safety and Security*. https://doi.org/10.1080/19439962.2018.1504262.

Westreich, D., Cole, S. R., Funk, M. J., Brookhart, M. A., & Stürmer, T. (2011). The role of the c-statistic in variable selection for propensity score models. *Pharmacoepidemiology Drug Safety, 20*(3), 317–320.

Wood, J., & Donnell, E. T. (2016). Safety evaluation of continuous green T intersections: A propensity scores-genetic matching-potential outcomes approach. *Accident Analysis and Prevention, 93*, 1–13.

Wood, J. S., & Donnell, E. T. (2017). Causal inference framework for generalizable safety effect estimates. *Accident Analysis and Prevention, 104*, 74–87.

Wood, J. S., Donnell, E. T., & Porter, R. J. (2015). Comparison of safety effect estimates obtained from empirical Bayes before–after study, propensity scores-potential outcomes framework, and regression model with cross-sectional data. *Accident Analysis and Prevention, 75*, 144–154.

World Health organization (WHO) (2018). Global status report on road safety 2018 Retrieved from https://www.who.int/violence injury prevention/road safety status/2018/en.

Haojie Li is currently an associate professor at School of Transportation, Southeast University. He holds a Ph.D. degree from Imperial College London. He used to work as a research associate at Railway and Transport Strategy Centre, within the Centre for Transport Studies at Imperial College London. His researches focus on risk and road accident analysis, statistical modeling for causal inference, and performance modelling and benchmarking.

Manman Zhu is currently a postgraduate student at School of Transportation, Southeast University, supervised by Dr. Haojie Li. She received a bachelor's degree in transportation engineering from Nanjing Agriculture University. Her research interests include traffic safety, public transportation systems, and statistical modeling for causal inference.

Daniel J. Graham is professor of Statistical Modelling in the Department of Civil and Environmental Engineering at Imperial College London (ICL) and Director of the ICL Transport Strategy Centre (TSC). He holds doctoral degrees from the Department of Mathematics at Imperial and from the London School of Economics. His research group works on mathematical and statistical models with applications in transportation, including: statistical modelling for performance analytics, causal inference methods and applications.

Gang Ren is currently a professor in transportation engineering and the director of the Major Projects and Collaborative Innovation Office of the university. His research interests include traffic behavior and safety analysis, emergency traffic management, and active transport optimization. He is the principal investigators of more than 10 research projects including three grants from the National Science Foundation of China (NSFC), and he has published more than 60 research papers in academic journals and conferences.

Journal of Safety Research 76 (2021) 327–331

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Evaluation of the National Electronic Injury Surveillance System – All injury program's self-directed violence data, United States, 2018 ☆

Daniel C. Ehlman [a,b,*], Tadesse Haileyesus [a], Robin Lee [a], Michael F. Ballesteros [a], Ellen Yard [a]

[a] Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, Division of Injury Prevention, 4770 Buford Highway, Atlanta, GA 30341, United States
[b] Centers for Disease Control and Prevention, Center for Surveillance, Epidemiology, and Laboratory Services, Epidemic Intelligence Service, 2400 Century Center, Atlanta, GA 30345, United States

ARTICLE INFO

Article history:
Received 1 December 2020
Accepted 2 December 2020
Available online 17 December 2020

Keywords:
Suicide
Self-harm
Self-directed violence
Surveillance
NEISS

ABSTRACT

Introduction: National estimates for nonfatal self-directed violence (SDV) presenting at EDs are calculated from the National Electronic Injury Surveillance System – All Injury Program (NEISS–AIP). In 2005, the Centers for Disease Control and Prevention and Consumer Product Safety Commission added several questions on patient characteristics and event circumstances for all intentional, nonfatal SDV captured in NEISS–AIP. In this study, we evaluated these additional questions along with the parent NEISS–AIP, which together is referred to as NEISS–AIP SDV for study purposes. *Methods:* We used a mixed methods design to evaluate the NEISS–AIP SDV as a surveillance system through an assessment of key system attributes. We reviewed data entry forms, the coding manual, and training materials to understand how the system functions. To identify strengths and weaknesses, we interviewed multiple key informants. Finally, we analyzed the NEISS–AIP SDV data from 2018—the most recent data year available—to assess data quality by examining the completeness of variables. *Results:* National estimates of SDV are calculated from NEISS–AIP SDV. Quality control activities suggest more than 99% of the cause and intent variables were coded consistently with the open text field that captures the medical chart narrative. Many SDV variables have open-ended response options, making them difficult to efficiently analyze. *Conclusions:* NEISS–AIP SDV provides the opportunity to describe systematically collected risk factors and characteristics associated with nonfatal SDV that are not regularly available through other data sources. With some modifications to data fields and yearly analysis of the additional SDV questions, NEISS–AIP SDV can be a valuable tool for informing suicide prevention. *Practical Applications:* NEISS-AIP may consider updating the SDV questions and responses and analyzing SDV data on a regular basis. Findings from analyses of the SDV data may lead to improvements in ED care.

1. Introduction

Every year in the United States nearly 50,000 people die due to suicide and nearly 500,000 present in emergency departments (EDs) for nonfatal self-directed violence (SDV) (also referred to as self-harm or self-inflicted injuries; Centers for Disease Control and Prevention National Centers for Injury Prevention and Control, 2020). This public health problem is worsening, as the age-adjusted rate of suicides and nonfatal SDV increased by 33% and 40%, respectively, between 2001 and 2018 (Centers for Disease Control and Prevention National Centers for Injury Prevention and Control, 2020). In 2018, suicide was the tenth leading cause of death in the United States (Centers for Disease Control and Prevention National Centers for Injury Prevention and Control, 2020). The direct and indirect costs for suicides and suicide attempts in the United States was estimated at $93.5 billion in 2013 (Shepard et al., 2016).

National estimates for nonfatal SDV presenting at EDs are calculated from the National Electronic Injury Surveillance System – All Injury Program (NEISS–AIP). NEISS–AIP is a collaboration between the U.S. Consumer Product Safety Commission (CPSC) and the U.S. Centers for Disease Control and Prevention (CDC) with the purpose

☆ **Special Report from the CDC:** *The Journal of Safety Research has partnered with the Office of the Associate Director for Science, Division of Injury Prevention, National Center for Injury Prevention and Control at the CDC in Atlanta, Georgia, USA, to briefly report on some of the latest findings in the research community. This report is the 64th in a series of "From the CDC" articles on injury prevention.*

* Corresponding author at: Applied Science Branch, Division of Injury Prevention, National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, 4770 Buford Highway NE, Mailstop S106-9, Atlanta, GA 30341, United States.

E-mail addresses: dehlman@cdc.gov (D.C. Ehlman), boq2@cdc.gov (T. Haileyesus), rpl5@cdc.gov (R. Lee), zzb0@cdc.gov (M.F. Ballesteros), igf8@cdc.gov (E. Yard).

https://doi.org/10.1016/j.jsr.2020.12.002
0022-4375/

of tracking first-time, nonfatal injury-related ED visits on all types and causes of injuries; deaths are excluded from the surveillance system. NEISS–AIP is a nationally-representative sample of 24-hour EDs with at least six beds. NEISS–AIP included 66 EDs when it started in 2000 and has decreased over time as more hospitals have dropped out than have been replaced; the 2018 sample included 59 hospitals.

CPSC and CDC train hospital coders to review ED medical records and abstract the necessary information on all injuries. CPSC manages and cleans the database with support from CDC. Select NEISS–AIP variables are available for querying through the WISQARS™ website (https://www.cdc.gov/injury/wisqars/index.html) within one year. Within 3–4 years the public version of the NEISS–AIP dataset can be accessed free through the Inter-University Consortium for Political and Social Research website (https://www.icpsr.umich.edu/web/ICPSR/series/198/studies) (Fig. 1).

In 2005, CDC and CPSC added several questions on patient characteristics and event circumstances for all intentional, nonfatal SDV cases captured in NEISS–AIP. The additional questions along with the parent NEISS–AIP constitute the NEISS–AIP SDV surveillance system for purposes of this study.

While aspects of NEISS–AIP and its other special studies have been evaluated in the past (Davis, Annest, Powell, & Mercy, 1996; Jhung et al., 2007; Thompson, Wheeler, Shi, Smith, & Xiang, 2014), NEISS–AIP SDV data have not been evaluated, so little is known about their usefulness. In this project, we evaluated NEISS–AIP SDV in terms of overall quality and utility.

2. Methods

CDC defines a surveillance system as "the ongoing, systematic collection, analysis, interpretation, and dissemination of data regarding a health-related event." These data are then used to inform prevention efforts "to reduce morbidity and mortality and to improve health" (Buehler, 1998; German, Horan, Lee, Milstein, & Pertowski, 2001; Teutsch & Thacker, 1995; Thacker, 2000). We used a mixed methods design to evaluate the NEISS–AIP SDV as a surveillance system through an assessment of 10 system attributes: usefulness, simplicity, flexibility, data quality, acceptability, sensitivity, predictive value positive, representativeness, timeliness, and stability (Table 1).

Table 1
Definitions of surveillance system attributes (German et al., 2001).

Attribute	Definition
Usefulness	A public health surveillance system is useful if it contributes to the prevention and control of adverse health-related events, including an improved understanding of the public health implications of such events
Simplicity	The simplicity of a public health surveillance system refers to both its structure and ease of operation. Surveillance systems should be as simple as possible while still meeting their objectives
Flexibility	A flexible public health surveillance system can adapt to changing information needs or operating conditions with little additional time, personnel, or allocated funds
Data Quality	Data quality reflects the completeness and validity of the data recorded in the public health surveillance system.
Acceptability	Acceptability reflects the willingness of persons and organizations to participate in the surveillance system.
Sensitivity	Sensitivity refers to the proportion of cases of a disease (or other health-related event) detected by the surveillance system (Weinstein & Fineberg)
Predictive Value Positive	Predictive value positive (PVP) is the proportion of reported cases that actually have the health-related event under surveillance (Weinstein & Fineberg)
Representativeness	A public health surveillance system that is representative accurately describes the occurrence of a health-related event over time and its distribution in the population by place and person
Timeliness	Timeliness reflects the speed between steps in a public health surveillance system
Stability	Stability refers to the reliability (i.e., the ability to collect, manage, and provide data properly without failure) and availability (i.e., the ability to be operational when it is needed) of the public health surveillance system

We reviewed data entry forms, the coding manual, and training materials to understand how the system functions. To identify strengths and weaknesses of the surveillance system attributes, we interviewed multiple key informants, including CDC users, CPSC managers, and hospital and quality assurance coders of the data. Finally, we analyzed the NEISS–AIP SDV data from 2018—

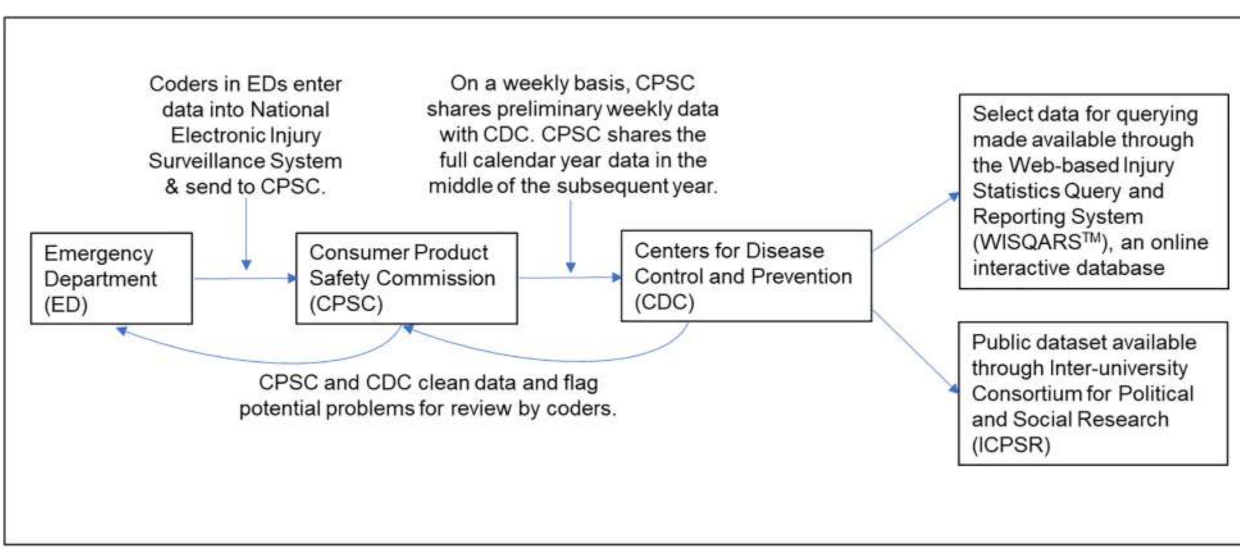


Fig. 1. National Electronic Injury Surveillance System – All Injury Program data flow.

the most recent data year available—to assess data quality by examining the completeness of variables; specific variables included time of arrival at the ED, patient self-reported SDV intent (e.g., intent to die, intent to harm oneself, intent to escape), staff description/diagnosis, patient SDV risk factors (e.g., previous episodes of self-harm, depression, bipolar disorder, anxiety), use of alcohol at time of injury, use of recreational drugs at time of injury, substances used (if poisoning), and final disposition (if admitted or transferred).

3. Results

3.1. Attributes

NEISS–AIP SDV allows for the calculation of national estimates of SDV. In addition, the system captures SDV-related variables that are not regularly available through other data sources, such as the Healthcare Cost and Utilization Product – Nationwide Emergency Department Sample. Compared to surveillance systems that rely on administrative codes alone, this system relies on medical record review, and, as such, might capture more cases. One study found that SDV-related administrative codes are frequently not recorded because, in part, they tend to not be billable; as a result, SDV events would be undercounted even though often there is enough information in the medical record to identify the SDV (Stanley et al., 2018). Sensitivity and predictive value positive were difficult to assess due to a lack of a gold standard for comparison. However, quality control activities suggest more than 99% of the cause and intent variables were coded consistently with the open text field that captures the medical chart narrative. In addition, hospital reporting to CPSC is timely as it occurs within a week of the ED visit, but data are not usually analyzed until after the calendar year's data have been cleaned and final weights have been assigned. Cleaning is completed about a year after data collection, which limits the ability to identify real-time changes in SDV-related trends. Findings from other system attributes can be found in Table 2.

3.2. Data quality (Completeness)

In 2018, NEISS–AIP SDV recorded 8,752 unweighted cases treated in EDs for nonfatal SDV injuries. Some variables (e.g., sex, age) do not have missing or unknown values but others (e.g., race, location where injury occurred, use of alcohol, use of recreational drugs, blood alcohol concentration (BAC)) have unknown values for more than 20% of observations (Table 3). Some variables only offer open-ended responses (e.g., BAC, poisoning substances and their respective quantities). A few variables (e.g., patient risk factors, BAC, poisoning substances and their respective quantities, patient disposition at ED discharge) have large numbers of missing data because responses are not required.

4. Discussion

NEISS–AIP SDV provides the opportunity to describe systematically collected risk factors and characteristics associated with non-

Table 2
Findings from the evaluation of attributes of the National Electronic Injury Surveillance System – All Injury Program Self-directed Violence surveillance system.

Attribute	Strengths	Weaknesses
Usefulness	• Provides annual national SDV estimates. • Captures SDV-related variables (e.g., risk factors, poisoning substances) that are not regularly available through other data sources.	• Open-ended responses used for many SDV-related variables make analysis time-consuming.
Simplicity	• Relatively few people involved in the data entry (generally just 1–2 coders per hospital). • Overall system data flow appears straightforward.	• NEISS–AIP (and by extension NEISS–AIP SDV) is not integrated with other data systems. • Data are manually abstracted from medical records.
Flexibility	• Relatively simple to add additional questions and responses onto existing NEISS–AIP structure, as evidenced by multiple additions that became effective in 2019.	• Changes require approval by multiple departments at CPSC, updates to the electronic data abstraction forms and new training for staff, all of which require funds as well.
Data Quality	• <1% of injuries had contradictory information in other dataset fields.	• Certain variables have many observations with no information because it is unknown or missing from the ED medical records.
Acceptability	• Hospital data are submitted complete and within seven days of ED visit, on average.	• ~1–3 hospitals drop out per year and need to be replaced (decline from 66 hospitals in 2000 to 59 in 2018).
Sensitivity	• Not possible to assess because medical records were not available. • Cases are captured through medical record review, which might identify more cases than would be captured by administrative codes alone.	
Predictive Value Positive	• Not possible to assess because medical records were not available. • <1% of injuries classified as SDV had contradictory information in other dataset fields.	
Representativeness	• Representative of U.S. hospitals having 24-hour EDs and a minimum of 6 inpatient beds that serve the general population (excludes Department of Veterans Affairs hospitals and special-purpose hospitals (e.g., correctional facilities, psychiatric-only hospitals)).	• Does not capture patients who do not seek medical treatment, those treated in physicians' offices or urgent care facilities and those treated in hospitals excluded from the sampling frame.
Timeliness	• Hospital data are submitted within seven days of ED visit, on average.	• Weighted data cannot be queried until approximately 1 year after data collection. • Public dataset not available until 3–4 years after data collection.
Stability	• CPSC's data reporting system is reliable. • CDC's WISQARS™ is reliably available for querying.	• Laptop problems and turn-over of hospital coders can lead to data entry delays.

CPSC = Consumer Product Safety Commission.
ED = Emergency department.
NEISS–AIP = National Electronic Injury Surveillance System – All Injury Program.
SDV = Self-directed violence.

Table 3
Description, type and evaluation findings of select 2018 National Electronic Injury Surveillance System – All Injury Program Self-directed Violence surveillance variables (8752 observations).

Variable Description	Variable Type	Findings
Age (in years)	Numeric	100% of observations have age.
Sex	Multiple Choice	100% of observations have sex.
Race	Multiple Choice	6572 (75%) of observations have race.
Location where injury occurred	Multiple Choice	5556 (63%) of observations have location.
Time of arrival to ED	Numeric	8708 (99%) of observations have time of arrival.
How did the patient describe his/her intent to the staff, other people, or in a (suicide) note?	Multiple Choice	100% of observations have patient-described intent.
Other description of intent	Open-ended	501 (6%) of observations have "other" descriptions
How did the staff describe or diagnose the injury event (at the time of discharge)?	Multiple Choice	100% of observations have staff description of injury.
Other staff description or diagnosis	Open-ended	837 (10%) of observations have "other" descriptions/diagnoses.
Depression	Checkbox	5379 (62%) of observations have this risk factor.
One or more previous episodes of self-harm	Checkbox	3235 (37%) of observations have this risk factor.
Anxiety, panic attacks, post-traumatic stress disorder	Checkbox	2048 (23%) of observations have this risk factor.
History of other substance(s) abuse	Checkbox	1091 (13%) of observations have this risk factor.
Other psychological/psychiatric problem, e.g., schizophrenia	Checkbox	1000 (11%) of observations have this risk factor.
Bipolar disorder	Checkbox	786 (9%) of observations have this risk factor.
History of alcohol abuse	Checkbox	697 (8%) of observations have this risk factor.
Borderline personality disorder	Checkbox	199 (2%) of observations have this risk factor.
Other specified risk factor(s) (e.g., argument with loved one, abuse or neglect, death of a loved one, illness, money or legal problems	Checkbox	3254 (37%) of observations have this risk factor.
Please specify the other risk	Open-ended	3254 (37%) of observations have a specific "other" risk factor.
Was alcohol used by the patient at the time of the injury event?	Multiple Choice	6753 (77%) of observations have information related to alcohol use.
Blood alcohol concentration (BAC) level	Open-ended	2042 (23%) of observations have BAC levels.
Were recreational drugs (e.g., cocaine, heroin, marijuana, ecstasy) used by the patient at the time of the injury event?	Multiple Choice	6580 (75%) of observations have information related to recreational drug use.
If the self-harm method was poisoning, please record up to four medications, drugs or substances taken by the patient. (4 "Substance" variables)	Open-ended	6123 (70%) of observations have "Substance 1". 1881 (22%) of observations have "Substance 2". 709 (8%) of observations have "Substance 3". 283 (3%) of observations have "Substance 4".
Amount substance taken (4 "Amount" variables)	Open-ended	6123 (70%) of observations have "Amount 1" (pertaining to "Substance 1"). 1881 (22%) of observations have "Amount 2" (pertaining to "Substance 2"). 709 (8%) of observations have "Amount 3" (pertaining to "Substance 3"). 283 (3%) of observations have "Amount 4" (pertaining to "Substance 4").
If the patient was admitted or transferred, please specify where s/he went	Multiple Choice	6453 (74%) of observations have information on patient disposition.

fatal SDV that are not regularly available through other data sources, which, in turn, would be useful for prevention purposes. While this surveillance system has the potential to be useful, this evaluation suggests that there are challenges with many of its system attributes.

The NEISS–AIP SDV surveillance system attributes of simplicity and stability benefit from being a part of the larger NEISS–AIP surveillance system. Another system strength is the focus on medical record review to capture cases, which is likely more sensitive than if the system relied only on administrative codes. Despite these strengths, NEISS–AIP has its limitations, particularly because it is currently reliant on human resources to manually abstract information from ED medical records and then enter the data into the NEISS–AIP data collection system. CPSC is exploring machine learning to help automate data abstraction from electronic medical records, but currently NEISS–AIP is not integrated with other data systems like electronic medical records.

NEISS–AIP SDV has aspects that are timely, including data reporting from the hospitals to CPSC (within a week of the ED visit) and feedback from CPSC and CDC to hospitals flagging certain errors (within about a week). However, it takes nearly a year after the end of the calendar year for the weighted data to be available for internal use at CDC and for select variables to be available to the general public through WISQARS™. Historically, the publicly available data set for NEISS–AIP was only available after 3–4 years. CDC is in the process of expediting the release of the public dataset that will allow for more timely analysis of the NEISS–AIP data by public health partners. The SDV data captured from the additional questions added in 2005 have not been and currently are not available to the public as we continue to evaluate their utility.

The usefulness of the NEISS–AIP SDV data requires further consideration due to a couple of system challenges. First, medical records that are incomplete or that do not require the same fields as NEISS–AIP SDV leads to unknown values being entered into NEISS–AIP SDV. For example, race is not always included in hospital ED records, which results in this variable frequently being missing and thus limiting the ability to look at associations between SDV and race.

In addition, many of the SDV-specific variables (e.g., patient self-reported intent, staff diagnosis, patient risk factors, BAC, poisoning substances and their respective quantities) use open text fields for large proportions of responses, making data entry time-consuming and data analysis difficult and inefficient. These variables should be examined to determine which questions and responses can be modified to reduce the amount of open-ended responses. For example, the poisoning substances variables could be displayed in a drop-down list by drug/substance class.

In addition to reducing the number of open-ended responses, there could be a review of which NEISS-AIP SDV variables should be maintained—either as is or with modifications—and which could be dropped because they are no longer relevant or because the data can be obtained elsewhere. For example, the NEISS–AIP SDV variable that captures staff description or diagnosis of the SDV injury is important to have a clinical assessment of self-harm intent but may need to be updated to avoid outdated terminology (e.g., suicide gesture) that has been in place since 2005. In addition, in 2019, NEISS–AIP added new variables to its core and modified the medical narrative field to collect more information regarding alcohol use, perhaps allowing for elimination of other alcohol use-related variables.

This evaluation was subject to at least two limitations. First, data were collected primarily through review of manuals and interviews with CPSC and CDC stakeholders. A standardized survey of all members of the NEISS–AIP surveillance team and a systematic review of medical records to validate the NEISS–AIP SDV data would have made this evaluation more robust. However, this was not possible due to time, financial, and planning constraints. Second, sensitivity and predictive value positive were difficult to assess due to a lack of a gold standard.

In summary, NEISS–AIP SDV is a unique surveillance system based on medical record review that collects SDV-related risk factors and characteristics that are not collected in other data sources. With some modifications to data fields and yearly analysis of the additional SDV questions, NEISS–AIP SDV can be a valuable tool for informing suicide prevention.

5. Funding source

This evaluation did not receive any specific grant from funding agencies in the commercial or not-for-profit sectors.

Acknowledgements

We would like to thank Tom Schroeder, Michelle White and Mary Cowhig at the Consumer Product Safety Commission, as well as the interviewed NEISS–AIP data entry and quality control staff, for their contributions to helping us better understand the surveillance system and its strengths and weaknesses.

Declaration of interest

None.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

References

Buehler, J. W. (1998). Surveillance. In K. J. Rothman & S. Greenland (Eds.), *Modern epidemiology* (2nd ed. Philadelphia, PA: Lippencott-Raven.

Centers for Disease Control and Prevention National Centers for Injury Prevention and & Control, (2020). Web-based Injury Statistics Query and Reporting System (WISQARS) [online] Retrieved Oct 16, 2020, from www.cdc.gov/injury/wisqars.

Davis, Yvette, Annest, Joseph L., Powell, Kenneth E., & Mercy, James A. (1996). An evaluation of the national electronic injury surveillance system for use in monitoring nonfatal firearm injuries and obtaining national estimates. *Journal of Safety Research, 27*(2), 83–91. https://doi.org/10.1016/0022-4375(96)00002-3.

German, R. R., Horan, J. M., Lee, L. M., Milstein, B., & Pertowski, C. A. (2001). Updated guidelines for evaluating public health surveillance systems; recommendations from the Guidelines Working Group. *MMWR Morb Mortal Wkly Rep, 50*(RR-13).

Jhung, Michael A., Budnitz, Daniel S., Mendelsohn, Aaron B., Weidenbach, Kelly N., Nelson, Theresa D., & Pollock, Daniel A. (2007). Evaluation and Overview of the National Electronic Injury Surveillance System-Cooperative Adverse Drug Event Surveillance Project (NEISS-CADES):. *Medical Care, 45*(Suppl 2), S96–S102. https://doi.org/10.1097/MLR.0b013e318041f737.

Shepard, Donald S., Gurewich, Deborah, Lwin, Aung K., Reed, Gerald A., Jr, & Silverman, Morton M. (2016). Suicide and Suicidal Attempts in the United States: Costs and Policy Implications. *Suicide Life Threat Behav, 46*(3), 352–362. https://doi.org/10.1111/sltb.12225.

Stanley, Barbara, Currier, Glenn W., Chesin, Megan, Chaudhury, Sadia, Jager-Hyman, Shari, Gafalvy, Hanga, & Brown, Gregory K. (2018). Suicidal Behavior and Non-Suicidal Self-Injury in Emergency Departments Underestimated by Administrative Claims Data. *Crisis, 39*(5), 318–325. https://doi.org/10.1027/0227-5910/a000499.

Teutsch, S. M., & Thacker, S. B. (1995). Planning a public health surveillance system. *Epidemiol Bull, 16*(1), 1–6.

Thacker, S. B. (2000). Historical development. In S. M. Teutsch & R. E. Churchill (Eds.), *Principles and practice of public health surveillance* (2nd ed. New York, NY: Oxford University Press.

Meghan C. Thompson Krista K. Wheeler Junxin Shi Gary A. Smith Huiyun Xiang Kim W. Carter An Evaluation of Comparability between NEISS and ICD-9-CM Injury Coding PLoS ONE 9 3 e92052 10.1371/journal.pone.0092052.s002.

Weinstein, M., & Fineberg, H. (1980). *Clinical Decision Analysis*. Philadelphia, PA: W. B. Saunders.

Daniel C. Ehlman ScD, MPH, is an Epidemic Intelligence Service (EIS) Officer assigned to the Division of Injury Prevention at CDC's National Center for Injury Prevention and Control. His research interests include suicide and self-directed violence prevention. Dr. Ehlman has a Doctor of Science degree in Global Health Systems and Development from the Tulane University School of Public Health and Tropical Medicine.

Tadesse Haileyesus MS, is a Mathematical Statistician working in the National Center for Injury Prevention and Control at the CDC. He received a master's degree in Statistics from the University of Georgia. He has authored and coauthored several scientific publications on a wide range of topics with a special focus on injury and violence prevention.

Robin Lee PhD, MPH, leads the Safety Promotion Team within CDC's National Center for Injury Prevention and Control. Dr. Lee and her team are actively engaged in studying the best ways to prevent injuries and creating real-world solutions to keep people safe, healthy, and independent. Topics of interest include drowning and injuries that impact older adults. Dr. Lee has a Master of Public Health and a Doctorate in Epidemiology from the State University of New York at Albany. She has authored and coauthored numerous presentations and scientific publications and has received awards for her public health and volunteer service.

Michael F. Ballesteros PhD, is the Deputy Associate Director for Science of the Division of Injury Prevention, National Center for Injury Prevention and Control, CDC. His research interests include injury surveillance systems, unintentional injuries, and global health. Dr. Ballesteros received a PhD in Epidemiology and is a graduate of CDC's Epidemic Intelligence Service (EIS) program.

Ellen Yard PhD, MPH, is an epidemiologist with the Suicide Prevention Team within CDC's National Center for Injury Prevention and Control. The Suicide Prevention Team strives towards a vision of no lives lost to suicide; this is carried out through the use of data, science, and partnerships to identify and implement effective suicide prevention strategies to foster healthy and resilient communities across the United States. Dr. Yard has a Doctorate in Epidemiology from The Ohio State University. She has authored and coauthored numerous presentations and scientific publications in injury prevention.

Journal of Safety Research 76 (2021) 101–117

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Exploring bicyclist injury severity in bicycle-vehicle crashes using latent class clustering analysis and partial proportional odds models

Zijing Lin [a], Wei (David) Fan [b,*]

[a] USDOT Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE), Department of Civil and Environmental Engineering, University of North Carolina at Charlotte, EPIC Building, Room 3366, 9201 University City Boulevard, Charlotte, NC 28223-0001, United States

[b] USDOT Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE), Department of Civil and Environmental Engineering, University of North Carolina at Charlotte, EPIC Building, Room 3261, 9201 University City Boulevard, Charlotte, NC 28223-0001, United States

ARTICLE INFO

Article history:
Received 11 June 2020
Received in revised form 17 July 2020
Accepted 23 November 2020
Available online 13 December 2020

Keywords:
Bicycle-involved crashes
Injury severity
Latent class analysis
Partial proportional odds models

ABSTRACT

Introduction: Bicyclists are more vulnerable compared to other road users. Therefore, it is critical to investigate the contributing factors to bicyclist injury severity to help provide better biking environment and improve biking safety. According to the data provided by National Highway Traffic Safety Administration (NHTSA), a total of 8,028 bicyclists were killed in bicycle-vehicle crashes from 2007 to 2017. The number of fatal bicyclists had increased rapidly by approximately 11.70% during the past 10 years (NHTSA, 2019). *Methods:* This paper conducts a latent class clustering analysis based on the police reported bicycle-vehicle crash data collected from 2007 to 2014 in North Carolina to identify the heterogeneity inherent in the crash data. First, the most appropriate number of clusters is determined in which each cluster has been characterized by the distribution of the featured variables. Then, partial proportional odds models are developed for each cluster to further analyze the impacts on bicyclist injury severity for specific crash patterns. *Results:* Marginal effects are calculated and used to evaluate and interpret the effect of each significant explanatory variable. The model results reveal that variables could have different influence on the bicyclist injury severity between clusters, and that some variables only have significant impacts on particular clusters. *Conclusions:* The results clearly indicate that it is essential to conduct latent class clustering analysis to investigate the impact of explanatory variables on bicyclist injury severity considering unobserved or latent features. In addition, the latent class clustering is found to be able to provide more accurate and insightful information on the bicyclist injury severity analysis. *Practical Applications:* In order to improve biking safety, regulations need to be established to prevent drinking and lights need to be provided since alcohol and lighting condition are significant factors in severe injuries according to the modeling results.

1. Introduction

Compared to other transportation modes, cycling is considered to be an environmentally friendly and sustainable means of travelling since it can help relieve the congestion pressure (Behnood & Mannering, 2017), reduce energy consumption and emissions (Pucher et al., 2011), and provide potential benefits in terms of environment, health, and society (Rojas-Rueda et al., 2011; Xia et al., 2013; Kelly et al., 2014; Macmillan et al., 2014; Götschi et al., 2016). Therefore, city planners and policy makers have been continuously encouraging and promoting cycling, and improving the bicycle facilities in order to construct a bicycle-friendly city and provide better cycling environment (Nabors et al., 2012).

As cycling has become more popular among citizens, especially for recreation and short-distance commuting trips (Klassen et al., 2014), there are certain issues that need to be addressed. One of the most critical concerns is cycling safety, which is highly associated with the fact that bicyclists are more vulnerable in comparison to other road users (Vanparijs, et al., 2015; Nilsson et al., 2017). From 2007 to 2017, there were 8,028 bicyclists killed in bicycle-motor vehicle crashes. The number of fatalities has increased by approximately 11.70% for the past 10 years. And in 2017, 50,000 bicyclists were injured accounting for 1.82% of total injuries in traffic crashes (NHTSA, 2019). In addition, bicyclists are found to be fatally injured with high probability especially in the United States (Pucher & Dijkstra 2003). Based on this situation, it is essential to identify and analyze the contributing factors to

* Corresponding author.
E-mail addresses: zlin4@uncc.edu (Z. Lin), wfan7@uncc.edu (Wei (David) Fan).

https://doi.org/10.1016/j.jsr.2020.11.012
0022-4375/

bicyclist injury severity resulting from bicycle-vehicle crashes using reasonable and appropriate modeling methods.

This paper aims to investigate the potential factors that significantly affect bicyclist injury severity using latent class clustering analysis and the partial proportional odds (PPO) model. The police reported bicycle-involved crash data collected from 2007 to 2014 in North Carolina were used in this study. Information including bicyclist, driver, vehicle, crash, roadway, temporal, and environmental characteristics was recorded in the database. Based on the crash data, latent class clustering analysis is conducted first to separate the whole data into homogenous segmentations, and PPO models are developed within each cluster to model the bicyclist injury severity. The contribution of this paper is the application of latent class clustering analysis to reduce the heterogeneity revealed in the bicycle-motor vehicle crash data and the method of combining it with a PPO model to explore the significant contributing factors to each type of crashes.

The remainder of this paper is structured as follows. Section 2 reviews previous studies on the relevant research topics. Section 3 describes the bicycle-motor vehicle crash data collected for this research study and the explanatory variables considered in bicyclist injury severity analysis. Section 4 explains the methodology used for this study including latent class clustering analysis and partial proportional odds model. Section 5 discusses the model results in detail. Finally, Section 6 summarizes this paper with a conclusion and provides recommendations for future work.

2. Literature review

Many researchers have conducted studies on the impact of different explanatory variables on bicyclist injury severity utilizing discrete choice models (Kim et al., 2007; Eluru et al., 2008; Yan et al., 2011; Kröyer, 2015; Behnood & Mannering, 2017; Chen et al., 2017; Robartes & Chen, 2017). However, some underlying conditions might exist in traffic crashes due to the heterogeneity, which results from the unobserved impact factors that cannot be reported or revealed from the collected data (Valent et al., 2002; Ulfarsson & Mannering, 2004; Pai & Saleh, 2007). This issue makes it difficult to analyze and evaluate the effects of significant factors on bicyclist injury severity resulting from such traffic crashes. In addition, the model bias cannot be neglected, which might lead to inaccurate conclusions (Shaheed & Gkritza, 2014; Mannering & Bhat, 2014). To overcome this problem, researchers have applied the segmentation method to resolve the heterogeneity issue by concentrating on specific crashes including crashes that occurred at different locations (Moore et al., 2011; Rash-ha Wahi et al., 2018; Lin & Fan, 2019a, 2019b), certain types of crashes (Decker et al., 2016), and different age groups (Gong & Fan, 2017; Li & Fan, 2019). However, the segmentation method mentioned above is usually based on the research need or the experience from the previous studies, which may not be able to guarantee the homogenous groups of the data (Depaire et al., 2008). Therefore, cluster analysis has been leveraged to separate the whole data and identify homogenous crash segmentations. Recently, latent class clustering analysis has been utilized (Depaire et al., 2008; Yasmin et al., 2014a, 2014b; Liu & Fan, 2020; Li & Fan, 2019) as the data clustering techniques to preprocess the distribution of data. While analyzing the injury severity considering the heterogeneity within each cluster, researchers adopted different models such as binary logit model (Sasidharan et al., 2015; Sivasankaran & Balasubramanian, 2020), multinomial logit model (Depaire et al., 2008; Sun et al., 2019), mixed logit model (Liu & Fan, 2018), and ordered probit model (Mohamed et al., 2013). To examine the unobserved heterogeneity underlying in the crash data, other data mining techniques including k-means clustering (Mohamed et al., 2013), decision tree and Bayesian networks (Prati et al., 2017), and classification and regression tree (Kashani & Mohaymany, 2011) are also utilized.

To model bicyclist injury severity, there are two categories of model structures that are ordered framework and unordered framework (Eluru, 2013). The basic discrete choice model within ordered framework is ordered logit model, which is employed for outcomes in ordinal nature (Mooradian et al., 2013). Based on this model, the partial proportional odds model relaxes the proportional odds assumption, which allows for variable coefficients across different levels. In addition to ordered framework models, the multinomial logit model and mixed logit models within unordered framework are also employed by researchers to conduct bicyclist injury severity analysis. However, the unordered models might neglect the inherent ordinal nature of injury severity.

By reviewing the previous relevant research studies, it can be concluded that there is a need to develop an innovative and combined method, which sequentially conducts the latent class clustering analysis and develops the partial proportional odds models, to model bicyclist injury severity in bicycle-vehicle crashes. The necessity of applying latent class clustering is to uncover the unobserved or latent features within the crash data, and to classify the bicycle-vehicle crashes into optimal homogenous groups for further analysis of heterogeneity between categorical segmentations. To explore the impact of various explanatory variables on bicyclist injury severity within each cluster, the partial proportional odds model can be developed in order to consider both the ordinal nature of injury severity levels and the limitation of proportional odds assumption associated with standard ordered logit models.

3. Data

The data utilized in this research study are the police reported bicycle-involved crash data collected from 2007 to 2014 in North Carolina, which record information including bicyclist injury severity, driver and bicyclist demographics, vehicle type and traveling speed, crash types and locations, roadway and environmental characteristics, and crash time, etc. The potential explanatory variables considered in this paper are carefully selected based on numerous previous studies as well as the availability of this dataset.

After data cleaning by removing the missing, unknown, and obviously incorrect data, a total of 4,012 bicycle-involved crash data are kept for descriptive analysis and model development. The bicyclist injury severity is categorized into five levels, which are no injury, possible injury, evident injury, disabling injury, and fatal injury, accounting for 8.85%, 38.96%, 43.62%, 5.76%, 2.81% total crashes, respectively. The same categorization method of injury severity levels can be found in research studies conducted by Eluru et al. (2008), and Liu et al. (2020).

Table 1 presents a descriptive analysis of the crash dataset, which contains the detailed information on the distribution of bicyclist injury severity levels and specific crashes. In addition, the potential factors obtained from the dataset are classified into eight categories to describe crashes from different aspects. The percentages of each injury severity level with different characteristics are presented in the table. The explanatory variables of the models are coded as dummy variables (0–1), which are listed in Table 1.

According to the data presented in Table 1, some potential variables that might be associated with severe injuries can be discovered. Alcohol usage for both bicyclists and drivers is a critical factor impacting bicyclist injury severity. Comparing the different effects on bicyclists and drivers, it can be seen that the percentage of fatal injuries associated with drivers in alcohol usage (16.00%) is higher

Table 1
Descriptive statistics of bicycle-motor vehicle crashes severity outcomes and explanatory variables.

Variable	Categories	Injury Severity					
	Description	Total	No injury	Possible injury	Evident injury	Disabling injury	Fatal injury
Bicycle-vehicle crashes		4012	8.85%	38.96%	43.62%	5.76%	2.81%
Cyclist characteristics							
Gender	Male	3426	9.28%	38.35%	43.70%	5.78%	2.89%
	Female	586	6.31%	42.49%	43.17%	5.63%	2.39%
Age	<16	726	9.37%	38.57%	45.04%	6.34%	0.69%
	16–24	860	9.88%	39.65%	42.56%	6.16%	1.74%
	25–54	1916	8.77%	39.25%	43.42%	5.22%	3.34%
	55+	510	6.67%	37.25%	44.12%	6.27%	5.69%
Alcohol usage	Yes	280	8.57%	30.36%	43.21%	10.00%	7.86%
	No	3732	8.87%	39.60%	43.65%	5.44%	2.44%
Driver characteristics							
Gender	Male	2194	8.48%	37.42%	44.80%	5.70%	3.60%
	Female	1818	9.30%	40.81%	42.19%	5.83%	1.87%
Age	<25	824	7.89%	37.74%	44.54%	5.83%	4.00%
	25–59	2418	9.47%	37.80%	44.13%	6.00%	2.61%
	60+	770	7.92%	43.90%	41.04%	4.94%	2.21%
Alcohol usage	Yes	75	4.00%	28.00%	38.67%	13.33%	16.00%
	No	3937	8.94%	39.17%	43.71%	5.61%	2.57%
Vehicle characteristics							
Veh_Type	Passenger Car	3015	9.59%	39.97%	42.82%	5.37%	2.26%
	Pickup	586	4.61%	38.05%	44.20%	8.19%	4.95%
	Van	221	4.98%	37.10%	50.23%	5.43%	2.26%
	Bus	25	24.00%	32.00%	36.00%	0.00%	8.00%
	Single Unit Truck	46	4.35%	21.74%	58.70%	6.52%	8.70%
	Motorcycle	26	15.38%	19.23%	61.54%	0.00%	3.85%
	Others	93	17.20%	32.26%	39.78%	6.45%	4.30%
Veh Speed	<20 mph	2052	10.87%	46.64%	40.11%	2.00%	0.39%
	20–30 mph	534	8.43%	33.71%	51.12%	5.62%	1.12%
	30–40 mph	650	6.62%	34.31%	46.92%	8.31%	3.85%
	40–50 mph	516	5.62%	26.94%	47.29%	13.18%	6.98%
	50–60 mph	251	5.58%	24.30%	41.04%	15.14%	13.94%
	60 + mph	9	11.11%	33.33%	22.22%	0.00%	33.33%
Crash characteristics							
Bike Direction	With traffic	2968	7.99%	35.14%	46.93%	6.70%	3.23%
	Facing traffic	1044	11.30%	49.81%	34.20%	3.07%	1.63%
Crash Type	Motorist overtaking bicyclist	839	5.96%	32.66%	45.77%	8.70%	6.91%
	Backing vehicle	34	23.53%	38.24%	38.24%	0.00%	0.00%
	Bicyclist failed to yield	610	8.69%	35.90%	43.77%	9.67%	1.97%
	Bicyclist Turn/Merge	367	10.08%	33.79%	43.60%	8.72%	3.81%
	Bicyclist overtaking motorist	55	16.36%	32.73%	49.09%	0.00%	1.82%
	Head-On	121	4.96%	32.23%	43.80%	9.92%	9.09%
	Motorist Failed to Yield	816	9.68%	52.94%	35.54%	1.47%	0.37%
	Motorist Turn/Merge	674	7.42%	36.05%	52.82%	3.12%	0.59%
	Crossing Paths	211	14.69%	46.45%	34.60%	3.79%	0.47%
	Parallel Paths	69	14.49%	30.43%	44.93%	7.25%	2.90%
	Other Crash Types	216	10.19%	37.96%	44.44%	4.17%	3.24%
Speeding	Yes	90	6.67%	22.22%	45.56%	12.22%	13.33%
	No	3922	8.90%	39.34%	43.57%	5.61%	2.58%
Rural/Urban	Urban	2700	9.78%	41.67%	42.59%	4.19%	1.78%
	Rural	1312	6.94%	33.38%	45.73%	8.99%	4.95%
Crash Location	Intersection	2329	9.32%	42.12%	42.81%	4.68%	1.07%
	Non-intersection	1683	8.20%	34.58%	44.74%	7.25%	5.23%
Roadway characteristics							
Road Geometry	Curve	230	5.65%	29.13%	52.17%	6.52%	6.52%
	Straight	3782	9.04%	39.56%	43.10%	5.71%	2.59%
Road Type	One-way	143	14.69%	46.15%	32.17%	4.90%	2.10%
	Two-way	3869	8.63%	38.69%	44.04%	5.79%	2.84%
Divided Road	Yes	801	9.11%	39.83%	43.07%	4.24%	3.75%
	No	3211	8.78%	38.74%	43.76%	6.14%	2.58%
Road Condition	wet, water, ice, snow, mud	323	8.05%	38.39%	44.89%	7.43%	1.24%
	dry	3689	8.92%	39.01%	43.51%	5.61%	2.95%
Traffic Control	Yes	2502	7.99%	41.49%	43.01%	5.44%	2.08%
	No	1510	10.26%	34.77%	44.64%	6.29%	4.04%
No. of Lanes	1	61	8.20%	40.98%	44.26%	6.56%	0.00%
	2	2395	8.89%	37.33%	44.51%	6.43%	2.84%
	3	316	10.13%	43.35%	41.14%	2.85%	2.53%
	4	611	8.84%	39.44%	42.88%	5.56%	3.27%
	5	394	7.11%	43.91%	40.61%	5.33%	3.05%
	6	127	8.66%	42.52%	42.52%	3.15%	3.15%
	7	38	7.89%	36.84%	50.00%	5.26%	0.00%

(*continued on next page*)

Table 1 (*continued*)

Variable	Categories	Injury Severity					
	Description	Total	No injury	Possible injury	Evident injury	Disabling injury	Fatal injury
	8	40	12.50%	35.00%	47.50%	2.50%	2.50%
	9+	30	13.33%	36.67%	43.33%	6.67%	0.00%
Land characteristics							
Work Zone	Yes	17	11.76%	29.41%	47.06%	11.76%	0.00%
	No	3995	8.84%	39.00%	43.60%	5.73%	2.83%
Temporal Characteristics							
Crash Time	0:00–5:59	137	7.30%	32.85%	39.42%	12.41%	8.03%
	6:00–9:59	547	7.31%	38.76%	46.98%	4.57%	2.38%
	10:00–14:59	1066	9.76%	42.78%	40.99%	4.13%	2.35%
	15:00–17:59	1093	9.15%	39.89%	43.37%	5.58%	2.01%
	18:00–23:59	1169	8.64%	35.41%	45.17%	7.19%	3.59%
Environmental Characteristics							
Weather	Clear	3326	8.81%	38.82%	44.11%	5.53%	2.74%
	Cloudy	511	8.61%	40.51%	40.90%	6.26%	3.72%
	Fog, smog, smoke	9	11.11%	33.33%	22.22%	0.00%	33.33%
	Rain	161	9.94%	36.02%	44.72%	9.32%	0.00%
	Snow	5	20.00%	80.00%	0.00%	0.00%	0.00%
Light Condition	Daylight	2970	8.99%	40.27%	43.80%	5.02%	1.92%
	Dusk or Dawn	176	6.82%	38.64%	44.32%	9.09%	1.14%
	Dark - lighted roadway	442	9.50%	37.33%	44.12%	6.79%	2.26%
	Dark - roadway not lighted	424	8.02%	31.60%	41.51%	8.49%	10.38%

than that for bicyclists (7.86%). In addition, vehicle speed is another variable that is related to bicyclist injury severity. Based on the bicyclist injury severity for different vehicle speeds shown in Table 1, the percentage of fatal injuries corresponding to each vehicle speed range increases with a higher vehicle speed. Similarly, speeding has a strong relationship with severe bicyclist injuries, as the percentage of fatal injuries resulting from speeding is 13.33%. Furthermore, environmental characteristics are the external causes to bicyclist injury severity. From the data presented in Table 1, the percentages of fatal injuries under adverse weather and dark not lighted roadway conditions are high.

4. Methodology

4.1. Latent class clustering (LCC)

Latent class clustering is a probability based cluster analysis approach (Depaire et al., 2008; Collins & Lanza, 2010), which has been widely used recently for traffic crash data segmentation in order to identify optimal homogenous groups. It is assumed that the whole crash dataset is divided into exclusive latent classes with each cluster being classified by an unobserved or latent categorical variable, which maximizes the homogeneity within each class and the heterogeneity between classes (Lanza & Rhoades, 2013; Sasidharan et al., 2015). Therefore, to uncover the unobserved and latent features underlying the bicycle-vehicle crash data, latent class clustering analysis is conducted. By applying this innovative method, the impact of various factors within the identified latent clusters on bicyclist injury severity can be investigated without omitting the potential heterogeneity between each segmentation. This method can be employed to conduct similar studies on injury severities of other road users. Compared to other clustering methods, latent class clustering does not require the number of clusters to be predetermined (Sun et al., 2019; Depaire et al., 2008). In addition, it allows different types of variables (e.g., numerical, categorical variables) (Sasidharan et al., 2015; Sun et al., 2019), and accepts different statistical criteria to determine the number of clusters (Sasidharan et al., 2015).

Let us consider a crash data sample where C latent clusters are assumed to be estimated based on K categorical items. Let $c = 1, 2, \ldots, C$ denote the latent class membership, and Y_i ($=Y_{i1}, \ldots, Y_{ik}$) represent crash i's responses to K categorical items in which Y_i is a categorical variable with possible values being $1, \ldots, r_k$. Let $I(y_k = r_k)$ be an indicator factor that equals to 1 when y_k equals to r_k, and 0 otherwise. Then, the probability function describing the response of crash is shown as follows:

$$P(Y_i = y) = \sum_{c=1}^{C} \gamma_c \prod_{k=1}^{K} \prod_{r_k=1}^{R_k} \rho_{k,r_k|c}^{I(y_k=r_k)} \tag{1}$$

where γ_c represents the probability of latent class membership for cluster c, and ρ denotes the item-response probability conditional on latent class membership. In this study, the latent class clustering analysis is conducted using the LCA procedure installed in SAS 9.4, which is developed by the Penn State Methodology Center (Lanza et al., 2007).

To determine the appropriate number of clusters, different number of clusters need to be tested by trying multiple models. Information criteria including Akaike Information Criterion (AIC), Bayesian Information Criteria (BIC), Consistent Akaike Information Criterion (CAIC), and entropy-based measures can be utilized to select the optimal number of clusters. The minimal values of AIC, BIC, CAIC indicate the best number of clusters. In addition, some researchers believe that BIC is better compared to AIC and CAIC for identifying the number of clusters (Biernacki & Govaert, 1999). However, it is also found that increasing the number of clusters might not always result in a minimum value, especially for large samples (e.g., traffic crash data) (Bijmolt et al., 2004). Therefore, the percentage reduction in BIC between tested models can be utilized (Sasidharan et al, 2015). As for the entropy measure, it is a weighted average of the posterior membership probability of an individual ranging from 0 to 1. Larger entropy values are associated with better latent class segmentation and 0.9 is suggested as a satisfied entropy value (McLachlan & Peel, 2000). In this paper, AIC, BIC, CAIC, and entropy measures are used for identifying the appropriate number of clusters.

4.2. Partial proportional odds model (PPO)

The partial proportional odds model is developed based on the ordered logit model. In the ordered logit model, the proportional odds (PO) assumption is subjected. It can be interpreted that the estimated parameters are restricted to be the same across all the alternatives. However, this assumption is unrealistic. To relax the assumption, the PPO model is developed.

The explanatory variables associated with each bicycle-vehicle crash are categorized into two groups. One contains parameters satisfying the PO assumption, which is presented as vector X_i; the other includes parameters that violate the PO assumption, which is shown as vector Z_i. The variables that violate the PO assumption are able to affect the response variables differently, while others remaining fixed parameters have the same effect across different levels. Thus, the PPO model with logit function is presented as follows (Peterson & Harrell, 1990):

$$P(Y_i \geq j) = \frac{\exp\left[\theta_j - \left(X_i^{'}\beta_j + Z_i^{'}\gamma_j\right)\right]}{1 + \exp\left[\theta_j - \left(X_i^{'}\beta_j + Z_i^{'}\gamma_j\right)\right]} \quad (2)$$

where j denotes the level of bicyclist injury severity and Y_i represents the crash injury resulting from bicycle-motor vehicle crash i, β and γ represent the coefficients that will be estimated, and θ_j demonstrates the threshold for jth cumulative logit.

To examine whether or not the explanatory variables violate the PO assumption, the Wald Chi-square tests are utilized during the model development (Wang & Abdel-Aty, 2008; Sasidharan & Menéndez, 2014). This procedure helps divide the explanatory variables into two groups that belong to either vector X_i or vector Z_i.

Before developing the PPO models, ordered logit models are built to help select the explanatory variables that will be considered later for the PPO models, and all parameters are assumed to violate the PO assumption as the base for PPO model development for each cluster. SAS 9.4 is used to conduct the PPO model estimation procedure. Since the sign of the estimated coefficient may not always explain the effect of explanatory variables on the bicyclist injury severity (Wooldridge, 2010; Washington et al., 2010), the marginal effects are applied for the PPO model result interpretation.

4.3. Marginal effect

To examine the impact of significant variables included in the partial proportional odds models on the likelihood of bicyclist injury severity, the marginal effect of each significant variable is calculated. Since all the variables in this research study are dummy-coded, it is not appropriate to apply the marginal effect equation for continuous variables (Yu et al., 2019). Therefore, the marginal effects of all the explanatory variables for each bicyclist i and severity level j are expressed as follows:

$$E_{X_{ijk}}^{P_{ij}} = P_{ij}(X_{ijk} = 1) - P_{ij}(X_{ijk} = 0) \quad (3)$$

where $E_{X_{ijk}}^{P_{ij}}$ represents the marginal effects of the k-th dummy variable X_{ijk}, $P_{ij}(X_{ijk} = 1)$ and $P_{ij}(X_{ijk} = 0)$ denote the probability when dummy variable X_{ijk} equals to 1 and 0, respectively.

The marginal effect of each significant explanatory variable for different bicyclist injury severity levels is calculated when the k-th dummy variable X_{ijk} varies from 0 to 1 based on the corresponding probabilities. Then, the average of marginal effects is computed for each parameter over all observations.

5. Results and discussions

5.1. Latent class clustering results

Several models are examined using the bicycle involved crash data including all the explanatory variables to identify the optimal number of clusters by testing from cluster 1 to cluster 10. Entropy value and information criteria including AIC, BIC, and CAIC are utilized to determine the most appropriate number of clusters and the variation of these values are shown in Fig. 1. Based on the values of AIC, BIC, and CAIC presented in this figure, all information criteria decrease with the increase of the number of clusters. However, no minimum value is reached for all the tested models. Hence, the reduction percentage of less than 1% in all three information criteria is applied to select the number of clusters for this study following the research conducted by Chang et al. (2019). From seven clusters afterwards, little improvement (<1%) of AIC, BIC, and CAIC is shown and the entropy has reached a high value of 0.97, which indicates a good separation between the clusters. Therefore, the bicycle-motor vehicle crashes are classified into seven clusters for further partial proportional odds analysis.

Following the research study conducted by Depaire et al. (2008), the final seven-cluster models can be described by the skewed feature distributions of each variable in each cluster, which can be found in Table 2. To characterize each cluster as a specific crash pattern, some featured variables are identified based on the variable distribution across clusters, and the variables with significantly different percentages are set to be bold in Table 2.

For cluster 1, 50.29% of the crashes occurred with estimated vehicle speed ranging from 20 to 30 mph. In addition, 55.32% of the crashes occurred when bicyclists failed to yield. Therefore, cluster 1 can be referred to as "crashes occurred with vehicle speed ranging from 20 to 30 mph when the bicyclist failed to yield." All the crashes in cluster 2 are caused by drivers aged from 25 to 59 with low estimated vehicle speed, which is less than 20 mph. One can define this cluster as "crashes caused by drivers from 25 to 59 years old with less than 20 mph vehicle speed." For cluster 3, 86.34% of the crashes occurred at non-intersection locations, and 61.8% occurred on dark roadway without light. Cluster 3 can be described as "crashes occurred on not lighted non-intersection roadways in dark." In cluster 4, only one variable is found to have significantly different distribution (63.81%) from that in other clusters, which is dark – lighted roadway. Hence, cluster 4 is named as "crashes occurred on lighted roadways in dark." Cluster 5 and cluster 6 overlap with cluster 3 on the same crash location, which is non-intersection location, but differ from it by the skewed feature of vehicle type (pickup) and lighting condition (daylight) respectively. It is noted that 65.43% of the crashes in cluster 5 are caused by pickups, and all the crashes in cluster 6 are occurred in daylight. Therefore, cluster 5 and cluster 6 can be referred to as "crashes caused by pickups at non-intersection locations" and "crashes occurred at non-intersection locations in daylight," respectively. Cluster 7 shares an overlapped variable (vehicle speed less than 20 mph) with cluster 2. All the crashes occurred with the estimated vehicle speed less than 20 mph, but none of them are caused by drivers aged from 25 to 59. Instead, 50.4% of the crashes are caused by old drivers (more than 60 years old). So, cluster 7 is described as "crashes caused by drivers elder than 60 years old with less than 20 mph vehicle speed." Finally, Table 3 summarizes the definition, key variables and the distribution of each cluster.

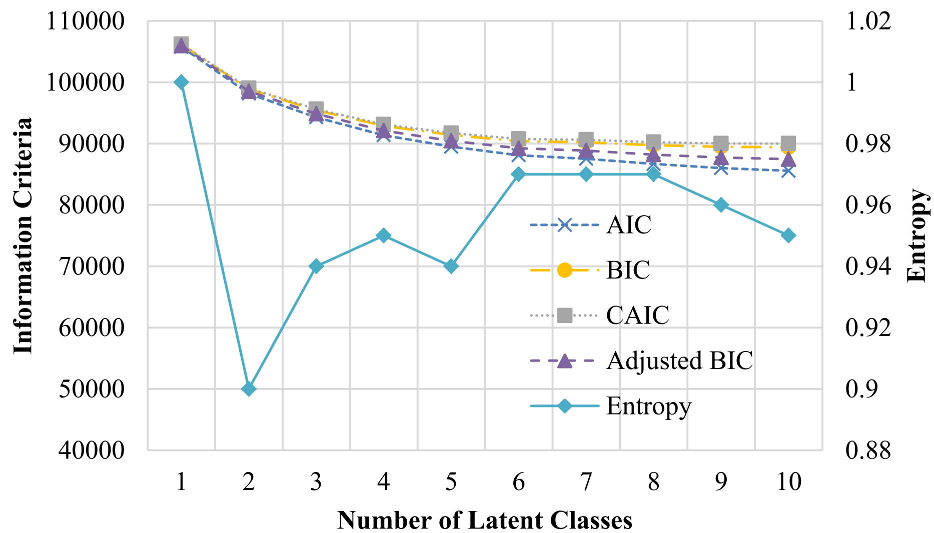


Fig. 1. Variation of entropy and information criteria for different number of clusters.

Table 2
Distribution of variables describing each cluster (bold).

Variable	Description	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
Bicycle-vehicle crashes		12.89%	24.08%	8.02%	13.81%	11.46%	14.21%	15.53%
Driver Age	25–59	61.9%	**100%**	62.11%	58.12%	63.7%	55.61%	0%
	60+	16.63%	0%	14.29%	16.06%	24.13%	21.75%	**50.4%**
Veh_Type	Pickup	12.19%	11.59%	0%	9.03%	**65.43%**	0%	9.63%
Veh Speed	<20 mph	0%	**100%**	4.35%	71.12%	5%	5.61%	**100%**
	20–30 mph	**50.29%**	0%	11.18%	12.82%	11.74%	19.82%	0%
Crash Type	Bicyclist failed to yield	**55.32%**	13.15%	1.86%	19.49%	0%	0.18%	13.16%
Crash Location	Non-intersection	2.13%	27.95%	**86.34%**	13.54%	**91.74%**	**84.56%**	23.27%
Light Condition	Daylight	99.81%	100%	0%	0%	64.13%	**100%**	100%
	Dark - lighted roadway	0%	0%	20.81%	**63.18%**	5.43%	0%	0%
	Dark - roadway not lighted	0%	0%	**61.8%**	19.31%	25.65%	0%	0%

Table 3
Definition and distribution of each cluster.

Cluster	Key Variables	Description	Percentage
Cluster 1	Vehicle speed 20–30 mph Bicyclist failed to yield	Crashes occurred with vehicle speed from 20 to 30 mph when the bicyclist failed to yield	12.89%
Cluster 2	Driver age 25–59 Vehicle speed less than 20 mph	Crashes caused by drivers from 25 to 59 years old with less than 20 mph vehicle speed	24.08%
Cluster 3	Crash location non-intersection Dark - roadway not lighted	Crashes occurred on not lighted non-intersection roadways in dark	8.02%
Cluster 4	Dark - lighted roadway	Crashes occurred on lighted roadways in dark	13.81%
Cluster 5	Vehicle type pickup Crash location non-intersection	Crashes caused by pickups at non-intersection locations	11.46%
Cluster 6	Crash location non-intersection Daylight	Crashes occurred at non-intersection locations in daylight	14.21%
Cluster 7	Driver age over 60 Vehicle speed less than 20 mph	Crashes caused by drivers elder than 60 years old with less than 20 mph vehicle speed	15.53%

5.2. Partial proportional odds model results

Based on the LCC results, PPO models are developed for each cluster. However, for cluster 4, cluster 6, and cluster 7, all the variables are found to pass the Wald Chi-square tests for the PO assumption, which make these three PPO models collapse into ordered logit models. To compare the restricted model developed based on the whole data and the sub-models developed based on latent class clustering, a PPO model is developed using the whole dataset, and a likelihood ratio test is conducted. According to the model estimation results, the log likelihood value at convergence for the restricted PPO model is −4530.08, while the sum of the log likelihood values at convergence for all seven sub-models is −4411.21. The value of χ^2 test statistics is 237.75 with 48 degrees of freedom, which indicates a better fitness for seven sub-models. Therefore, the PPO/ORL model results for each cluster and the whole data are presented in Tables 4a–4h.

As is mentioned in Section 4.2., the sign of the estimated parameters may not accurately reveal the effect of bicyclist injury severity, marginal effects need to be calculated for the interpretation of the variable impacts. The marginal effects of significant variables in each model for each cluster and the whole data are presented in Tables 5a–5h.

Comparing the significant factors for the whole dataset and separate clusters, three critical findings can be discussed. First, differences can be found between the significant explanatory variables for the whole dataset and the seven clusters. Some variables that do not significantly affect the bicyclist injury severity in the whole dataset are found to have significant impacts in specific clusters

Table 4a
PPO model for cluster 1 in bicyclist-vehicle crashes.

Cluster 1		All Levels		Fatal Injury		Disabling Injury		Evident Injury		Possible Injury	
Variable		Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.
Intercept				−3.0494	0.0010**	−1.1260	0.1405	0.0559	0.9474	0.5546	0.5538
Vehicle characteristics											
Veh Speed	20–30 mph			−3.1930	0.0025**	−2.6773	<0.0001**	−0.7646	0.2706	0.7874	0.3418
	30–40 mph			−5.4050	0.0004**	−2.5565	<0.0001**	−1.1038	0.1131	0.2602	0.7511
	40–50 mph			−4.0777	0.0120**	−1.3928	0.0352**	−0.3270	0.6606	2.2216	0.0828*
Crash characteristics											
Bike Direction	Facing traffic			1.5879	0.0837*	−0.3658	0.3682	−0.5812	0.0119**	−0.0941	0.8110
Crash Type	Bicyclist Turn/Merge	1.0534	0.0458**								
	Bicyclist overtaking motorist	−2.8403	0.0325**								
	Motorist Failed to Yield	−0.5971	0.0500*								
Roadway characteristics											
Road Geometry	Curve			4.1579	0.0011**	0.4645	0.4458	0.6454	0.2220	13.4334	0.9851
Road Type	Two-way	1.0816	0.0170**								
Model Performance Results											
Log Likelihood with Constant Only		−619.40									
Log Likelihood at Convergence		−577.81									
AIC		1213.63									

* Level of significance >90%.
** Level of significance >95%.

Table 4b
PPO model for cluster 2 in bicyclist-vehicle crashes.

Cluster 2		All Levels		Fatal Injury		Disabling Injury		Evident Injury		Possible Injury	
Variable		Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.
Intercept				−18.8777	0.9821	−4.2560	<0.0001**	−0.6332	0.1332	2.8783	<0.0001**
Cyclist characteristics											
Gender	Male			11.7300	0.9889	−0.1434	0.8007	−0.2588	0.1502	−1.6871	0.0003**
Vehicle characteristics											
Veh_Type	Passenger Car	0.8759	0.0089**								
	Pickup	1.1659	0.0020**								
	Van	0.9009	0.0284**								
	Single Unit Truck	2.1506	0.0052**								
	Motorcycle	2.4802	0.0436**								
Crash characteristics											
Bike Direction	Facing traffic			−13.0156	0.9877	−1.1769	0.0370**	−0.7558	<0.0001**	−0.2122	0.3103
Crash Type	Motorist Turn/Merge	0.3634	0.0309**								
Rural/Urban	Urban	−0.4694	0.0080**								
Roadway characteristics											
Road Geometry	Curve	0.7416	0.0384**								
Road Condition	wet, water, ice, snow, mud	0.8671	0.0028**								
Traffic Control	Yes	0.3975	0.0042**								
Temporal Characteristics											
Crash Time	0:00–5:59	1.9987	0.0377**								
	6:00–9:59	−0.3212	0.0635*								
Environmental Characteristics											
Weather	Fog, smog, smoke, rain, snow	−0.4284	0.0279**								
Model Performance Results											
Log Likelihood with Constant Only		−1029.52									
Log Likelihood at Convergence		−971.68									
AIC		1993.37									

* Level of significance >90%.
** Level of significance >95%.

(i.e., types of bicycle-vehicle crashes). This finding confirms the necessity and importance of conducting the latent class clustering analysis, which can provide the insight into the differences between various types of crashes and reveal the heterogeneity within the data. Second, the significant explanatory variables vary across the seven clusters, indicating that different types of crashes are affected by distinctive factors. Third, significant variables in the PPO model that are developed based on the whole dataset can still be found in the sub-models, which provides clear evidence that the sub-models can interpret the effects of bicyclist injury severity more exhaustively.

Differences of the impacts of significant variables identified across seven clusters will be discussed in detail in the following sections.

Table 4c
PPO model for cluster 3 in bicyclist-vehicle crashes.

Cluster 3		All Levels		Fatal Injury		Disabling Injury		Evident Injury		Possible Injury	
Variable		Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.
Intercept				−3.1950	<0.0001**	−1.9005	<0.0001**	0.3748	0.0197**	2.5665	<0.0001**
Driver characteristics											
Age	<25	0.7157	0.0039**								
Alcohol usage	Yes	1.2941	0.0104**								
Vehicle characteristics											
Veh Speed	50–60 mph			1.5335	0.0004**	1.0300	0.0026**	0.1141	0.7237	−0.1120	0.8345
Crash characteristics											
Bike Direction	Facing traffic			−1.3387	0.1420	0.4606	0.3477	−0.8320	0.0499**	−1.2660	0.0239**
Speeding	Yes	1.4298	0.0154**								
Roadway characteristics											
Road Condition	wet, water, ice, snow, mud	−0.8877	0.0296**								
Temporal Characteristics											
Crash Time	15:00–17:59			1.5150	0.1110	−1.0948	0.0808*	−0.6792	0.0783*	−1.0940	0.0366**
Environmental Characteristics											
Weather	Fog, smog, smoke, rain, snow	0.9195	0.0053**								
Model Performance Results											
Log Likelihood with Constant Only		−454.59									
Log Likelihood at Convergence		−971.68									
AIC		879.97									

* Level of significance >90%.
** Level of significance >95%.

Table 4d
ORL model for cluster 4 in bicyclist-vehicle crashes.

Cluster 4		All Levels		Fatal Injury		Disabling Injury		Evident Injury		Possible Injury	
Variable		Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.
Intercept				−3.3756	<0.0001**	−1.6385	<0.0001**	1.2704	<0.0001**	3.6950	<0.0001**
Cyclist characteristics											
Age	<16	−0.4630	0.0710*								
Vehicle characteristics											
Veh Speed	<20 mph	−1.3416	<0.0001**								
	20–30 mph	−0.8078	0.0106**								
Crash characteristics											
Crash Type	Bicyclist overtaking motorist	−1.4275	0.0723*								
	Motorist Failed to Yield	−0.6251	0.0008**								
	Parallel Paths	−1.7609	0.0029**								
Speeding	Yes	−1.4821	0.0455**								
Model Performance Results											
Log Likelihood with Constant Only		−637.81									
Log Likelihood at Convergence		−607.89									
AIC		1237.78									

* Level of significance >90%.
** Level of significance >95%.

5.2.1. Bicyclist characteristics

Bicyclist characteristic factors that have significant impacts on bicyclist injury severity include the gender of bicyclists, bicyclist age, and bicyclists under the influence of alcohol. Differences of impact variables exist between each cluster. For example, male bicyclists in cluster 2 are found to be more likely to have fatal injury or no injury with marginal effects being 0.0012 and 0.1079, respectively.

Young bicyclists (<16) have a lower probability of suffering from severe injuries (fatal injury, disabling injury, and evident injury) in cluster 4 (marginal effects −0.0042, −0.0167, and −0.0822) and cluster 5 (marginal effects −0.06, −0.0512, and −0.114). Similarly, bicyclists (16–24 and 25–54) in cluster 5 are less likely to sustain severe injury, including fatal injury, disabling injury, and evident injury (marginal effects −0.0476, −0.038, and −0.0697; and marginal effects −0.0748, −0.054, and −0.053). This result is in line with a previous research study conducted by Kaplan et al. (2014).

In cluster 7, bicyclists under the influence of alcohol have a higher likelihood to be severely injured (fatal injury, disabling injury, and evident injury). It can be noted that, although the estimated driving speed is relatively low, the injury severity level can be high. This result is probably valid because drinking alcohol might decrease the reaction speed of a bicyclist to an incident and therefore have a negative impact on his/her physical condition. In addition, the influence of elderly drivers cannot be neglected. Therefore, regulations can be made to prevent bicyclists from drinking alcohol while riding a bicycle, so that biking safety might be improved.

Table 4e
PPO model for cluster 5 in bicyclist-vehicle crashes.

Cluster 5		All Levels		Fatal Injury		Disabling Injury		Evident Injury		Possible Injury	
Variable		Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.
Intercept				−0.8829	0.0416**	−0.0170	0.9678	2.4620	<0.0001**	5.0348	<0.0001**
Cyclist characteristics											
Age	<16	−1.0458	0.0022**								
	16–24	−0.7379	0.0214**								
	25–54	−0.9294	0.0002**								
Driver characteristics											
Age	25–59	0.4041	0.0348**								
Vehicle characteristics											
Veh_Type	Single Unit Truck	0.8254	0.0665*								
Veh Speed	<20 mph	−1.1364	0.0076**								
	20–30 mph	−0.9011	0.0021**								
	30–40 mph	−0.5539	0.0112**								
Crash characteristics											
Speeding	Yes			1.6194	0.0046**	1.2554	0.0160**	0.9392	0.1619	−1.1144	0.1742
Roadway characteristics											
No. of Lanes	<=4	−0.7549	0.0247**								
Environmental Characteristics											
Light Condition	Daylight	−0.3785	0.0611*								
	Dark - lighted roadway	−0.8676	0.0487**								
Model Performance Results											
Log Likelihood with Constant Only		−584.12									
Log Likelihood at Convergence		−552.88									
AIC		1143.76									

* Level of significance >90%.
** Level of significance >95%.

Table 4f
ORL model for cluster 6 in bicyclist-vehicle crashes.

Cluster 6		All Levels		Fatal Injury		Disabling Injury		Evident Injury		Possible Injury	
Variable		Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.
Intercept				−0.4500	0.7383	0.8233	0.5412	3.2915	0.0151**	5.4747	<0.0001**
Vehicle characteristics											
Veh Speed	30–40 mph	0.4648	0.0262**								
	40–50 mph	0.7711	0.0004**								
	50–60 mph	1.0369	0.0002**								
Crash characteristics											
Crash Type	Bicyclist Turn/Merge	−0.4803	0.0074**								
	Head-On	1.2955	0.0005**								
Roadway characteristics											
Road Type	Two-way	−2.8927	0.0266**								
Traffic Control	Yes	0.3239	0.0514*								
No. of Lanes	<=4	−0.5629	0.0661*								
Environmental Characteristics											
Weather	Fog, smog, smoke, rain, snow	−0.3908	0.0882*								
Model Performance Results											
Log Likelihood with Constant Only		−718.46									
Log Likelihood at Convergence		−692.91									
AIC		1411.81									

* Level of significance >90%.
** Level of significance >95%.

5.2.2. Driver characteristics

Young drivers might increase the probability of severe injuries (fatal injury, disabling injury, and evident injury) suffered by bicyclists in crashes that occurred on not lighted, non-intersection locations in dark with marginal effects being 0.0673, 0.0513, and 0.027. Similar results can be found in cluster 5, where mid-aged drivers (25–59) are more likely to cause severe injuries including fatal, disabling, and evident injury in crashes involving pickups at non-intersection locations.

Unsurprisingly, drivers under the influence of alcohol can probably increase the likelihood of getting severe injuries such as fatal and disabling injuries (marginal effects 0.1501 and 0.0976) for crashes occurred on not lighted non-intersection locations in dark. This result is in line with previous research in (Kim et al., 2007; Moore et al., 2011). Policies can be implemented to prohibit drivers from drinking alcohol to avoid severe injuries to bicyclists in bicycle-vehicle crashes.

5.2.3. Vehicle characteristics

Several types of vehicles are found to have significant impacts on bicyclist injury severity, especially in cluster 2 and cluster 5, which are crashes caused by drivers from 25 to 59 years old with

Table 4g
ORL model for cluster 7 in bicyclist-vehicle crashes.

Cluster 7		All Levels		Fatal Injury		Disabling Injury		Evident Injury		Possible Injury	
Variable		Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.
Intercept				−4.6780	<0.0001**	−3.6478	<0.0001**	0.2588	0.1795	3.2116	<0.0001**
Cyclist characteristics											
Alcohol usage	Yes	2.0006	0.0115**								
Crash characteristics											
Bike Direction	Facing traffic	−0.3601	0.0435**								
Crash Type	Motorist Turn/Merge	0.5508	0.0044**								
	Crossing Paths	−1.4198	<0.0001**								
	Parallel Paths	−2.5955	0.0241**								
Roadway characteristics											
Road Geometry	Curve	0.8632	0.0394**								
Divided Road	Yes	−0.4608	0.0164**								
Temporal Characteristics											
Crash Time	10:00–14:59	−0.6423	0.0017**								
	15:00–17:59	−0.5938	0.0050**								
Model Performance Results											
Log Likelihood with Constant Only		−626.19									
Log Likelihood at Convergence		−589.05									
AIC		1204.11									

*Level of significance >90%.
** Level of significance >95%.

Table 4h
PPO model for the whole data in bicyclist-vehicle crashes.

Whole Data		All Levels		Fatal Injury		Disabling Injury		Evident Injury		Possible Injury	
Variable		Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.	Coef.	Sig.
Intercept				−3.3425	<0.0001**	−2.3343	<0.0001**	0.0064	0.9756	2.2166	<0.0001**
Cyclist characteristics											
Age	<16			−1.6649	0.0007**	−0.5829	0.0025**	−0.2271	0.0585*	−0.3235	0.0640*
	16–24			−0.9676	0.0011**	−0.2544	0.1355	−0.1984	0.0797*	−0.3341	0.0358**
	25–54	−0.2531	0.0085**								
Alcohol usage	Yes			0.5943	0.0192**	0.5355	0.0031**	0.2172	0.1061	−0.1600	0.4877
Driver characteristics											
Age	<25	0.2511	0.0092**								
	25–59	0.1332	0.0909*								
Alcohol usage	Yes			1.5997	<0.0001**	1.1884	<0.0001**	0.4384	0.0922*	0.6569	0.2690
Vehicle characteristics											
Veh Speed	<20 mph			−2.1048	<0.0001**	−1.5920	<0.0001**	−0.5245	<0.0001**	−0.3628	0.0196**
	20–30 mph			−1.0894	0.0107**	−0.8127	<0.0001**	−0.0567	0.6307	−0.1684	0.4075
	40–50 mph	0.3912	0.0005**								
	50–60 mph			1.1516	<0.0001**	0.9503	<0.0001**	0.4712	0.0034**	0.1895	0.5353
Veh_Type	Pickup			0.3012	0.1733	0.2610	0.0761*	0.1139	0.2343	0.7111	0.0006**
	Van	0.2453	0.0599*								
	Single Unit Truck	0.7171	0.0136**								
Crash characteristics											
Bike Direction	Facing traffic			−0.1770	0.5107	−0.4532	0.0092**	−0.4762	<0.0001**	−0.1654	0.2090
Speeding	Yes	0.3182	0.1275								
Crash Type	Bicyclist Failed to Yield			0.2613	0.4250	0.8624	<0.0001**	0.3064	0.0018**	0.1970	0.2342
	Head-On	0.7386	<0.0001**								
	Motorist Turn/Merge			−0.6942	0.1939	−0.1762	0.4417	0.4386	<0.0001**	0.3913	0.0210**
Roadway characteristics											
Road Geometry	Curve	0.4310	0.0012**								
Road Type	Two-way	0.2931	0.0772*								
Temporal Characteristics											
Crash Time	10:00–14:59	−0.1334	0.0535*								
Model Performance Results											
Log Likelihood with Constant Only		−4848.98									
Log Likelihood at Convergence		−4530.08									
AIC		9178.16									

* Level of significance >90%.
** Level of significance >95%.

less than 20 mph vehicle speed, and crashes caused by pickups at non-intersection locations. To be specific, passenger cars, pickups, vans, single unit trucks, and motorcycles are types of vehicles that might increase the probability of severe injuries (fatal, disabling, and evident injuries). Similar results can be found in Gårder (1994), Gårder et al. (1998), and Stone and Broughton (2003).

Table 5a
Average marginal effects for cluster 1.

Variable		Cluster1				
		F	D	E	P	N
Vehicle characteristics						
Veh Speed	20–30 mph	−0.1230	−0.1726	0.1349	0.2164	−0.0558
	30–40 mph	−0.1100	−0.1214	−0.0145	0.2641	−0.0182
	40–50 mph	−0.0470	−0.0463	0.0197	0.1546	−0.0811
Crash characteristics						
Bike Direction	Facing traffic	0.0385	−0.0687	−0.1042	0.1277	0.0067
Crash Type	Bicyclist Turn/Merge	0.0273	0.1042	0.0682	−0.1498	−0.0499
	Bicyclist overtaking motorist	−0.0228	−0.0786	−0.4150	0.0600	0.4564
	Motorist Failed to Yield	−0.0099	−0.0348	−0.0933	0.0878	0.0502
Roadway characteristics						
Road Geometry	Curve	0.2642	−0.2161	0.0842	−0.0501	−0.0822
Road Type	Two-way	0.0150	0.0530	0.1821	−0.1404	−0.1096

N - No Injury.
P - Possible Injury.
E - Evident Injury.
D - Disabling Injury.
F - Fatal Injury.

Table 5b
Average marginal effects for cluster 2.

Variable		Cluster2				
		F	D	E	P	N
Cyclist characteristics						
Gender	Male	0.0012	−0.0045	−0.0551	−0.0496	0.1079
Vehicle characteristics						
Veh_Type	Passenger Car	0.0008	0.0161	0.1621	−0.0732	−0.1058
	Pickup	0.0021	0.0380	0.2177	−0.1711	−0.0867
	Van	0.0014	0.0274	0.1614	−0.1212	−0.0690
	Single Unit Truck	0.0073	0.1203	0.2974	−0.3209	−0.1041
	Motorcycle	0.0105	0.1610	0.2910	−0.3550	−0.1075
Crash characteristics						
Bike Direction	Facing traffic	−0.0017	−0.0197	−0.1499	0.1499	0.0215
Crash Type	Motorist Turn/Merge	0.0004	0.0080	0.0741	−0.0486	−0.0339
Rural/Urban	Urban	−0.0006	−0.0111	−0.0954	0.0652	0.0418
Roadway characteristics						
Road Geometry	Curve	0.0011	0.0209	0.1472	−0.1113	−0.0578
Road Condition	wet, water, ice, snow, mud	0.0013	0.0253	0.1701	−0.1298	−0.0669
Traffic Control	Yes	0.0004	0.0076	0.0794	−0.0454	−0.0420
Temporal Characteristics						
Crash Time	0:00–5:59	0.0061	0.1049	0.2934	−0.3040	−0.1005
	6:00–9:59	−0.0003	−0.0060	−0.0638	0.0353	0.0348
Environmental Characteristics						
Weather	Fog, smog, smoke, rain, snow	−0.0004	−0.0079	−0.0842	0.0449	0.0476

N - No Injury.
P - Possible Injury.
E - Evident Injury.
D - Disabling Injury.
F - Fatal Injury.

Another significant impact factor of vehicle characteristic is the estimated vehicle speed, which is highly associated with the bicyclist injury severity. Low vehicle speed (less than 30 mph) is less likely to result in severe injuries including fatal, disabling, and evident injuries in cluster 4 and cluster 5 (i.e., crashes occurred on lighted roadways in dark and crashes caused by pickups at non-intersection locations, respectively). Interestingly, the estimated vehicle speed from 30 to 40 mph has different effects for cluster 5 and cluster 6. In cluster 5, this range of vehicle speed is less likely to cause severe injuries for crashes caused by pickups at non-intersection locations, while in cluster 6, the vehicle speed might increase the probability of severe injuries for crashes occurred at non-intersection locations in daylight. For vehicle speed from 40 to 60 mph, it is clear that the likelihood of suffering from severe injuries is increased for crashes occurring at non-intersection locations in daylight. It can be seen that even in daylight, driving fast may increase the probability of severe injuries of bicyclists resulting from crashes occurred at non-intersection locations. Therefore, speed limit is critical for providing safer cycling environment. When determining the speed limit for roadways that are popular among bicyclists, one needs to carefully consider avoiding setting the speed limit over 40 mph based on the model estimation results since high speed may increase the likelihood of severe injuries of bicyclists.

Table 5c
Average marginal effects for cluster 3.

Variable		Cluster3				
		F	D	E	P	N
Driver characteristics						
Age	<25	0.0673	0.0513	0.0270	−0.1040	−0.0415
Alcohol usage	Yes	0.1501	0.0976	−0.0241	−0.1675	−0.0560
Vehicle characteristics						
Veh Speed	50–60 mph	0.1708	0.0098	−0.1565	−0.0320	0.0079
Crash characteristics						
Bike Direction	Facing traffic	−0.0785	0.1541	−0.2622	0.0588	0.1277
Speeding	Yes	0.1744	0.1046	−0.0416	−0.1792	−0.0582
Roadway characteristics						
Road Condition	wet, water, ice, snow, mud	−0.0606	−0.0526	−0.0814	0.1155	0.0791
Temporal Characteristics						
Crash Time	15:00–17:59	0.1785	−0.3097	−0.0211	0.0497	0.1026
Environmental Characteristics						
Weather	Fog, smog, smoke, rain, snow	0.0890	0.0663	0.0247	−0.1273	−0.0527

N - No Injury.
P - Possible Injury.
E - Evident Injury.
D - Disabling Injury.
F - Fatal Injury.

Table 5d
Average marginal effects for cluster 4.

Variable		Cluster4				
		F	D	E	P	N
Cyclist characteristics						
Age	< 16	−0.0042	−0.0167	−0.0822	0.0563	0.0468
Vehicle characteristics						
Veh Speed	<20 mph	−0.0178	−0.0688	−0.2189	0.2107	0.0948
	20–30 mph	−0.0070	−0.0277	−0.1360	0.0798	0.0910
Crash characteristics						
Crash Type	Bicyclist overtaking motorist	−0.0082	−0.0344	−0.2374	0.0777	0.2023
	Motorist Failed to Yield	−0.0056	−0.0227	−0.1153	0.0840	0.0596
	Parallel Paths	−0.0092	−0.0385	−0.2786	0.0572	0.2692
Speeding	Yes	−0.0085	−0.0355	−0.2430	0.0745	0.2125

N - No Injury.
P - Possible Injury.
E - Evident Injury.
D - Disabling Injury.
F - Fatal Injury.

5.2.4. Crash characteristics

Biking direction is a significant impact factor that affects bicyclist injury severity. Since bicyclists will have the ability to prepare while biking facing traffic, the probability of suffering from severe injury severities is decreased, which is consistent with the model results revealed in cluster 2, cluster 3, and cluster 7 (i.e., crashes caused by mid-aged drivers with low speed, crashes occurred on not lighted non-intersection locations in dark, and crashes caused by elderly drivers with low speed). However, in cluster 1, bicyclists are more likely to sustain fatal injury (marginal effect 0.0385), which is related to the specific crash pattern (crashes occurred with vehicle speed from 20 to 30 mph when bicyclists failed to yield). This result indicates that severe injuries might occur due to the fault of bicyclists in bicycle-vehicle crashes. This is consistent with a research study conducted by Kim et al. (2007). Considering this model results, it is necessary to specify the right of way when there is a conflict between drivers and bicyclists. A yield sign will help remind drivers of yielding to bicyclists to reduce the probability of a collision. Furthermore, bicyclists need to wear reflective materials so as to be clearly seen by other road users.

Different crash types will affect bicyclist injury severity distinctively, and different effects may exist for specific crash patterns. In cluster 1, more severe injuries might be suffered by bicyclists when bicyclists turn or merge, while in cluster 6, the opposite results can be concluded that bicyclists are less likely to be severely injured in crashes occurred at non-intersection locations in daylight. That is probably associated with the crash locations (non-intersection) and the lighting condition (daylight). When bicyclists overtake motorist, the likelihood of bicyclists getting severe injuries (including fatal, disabling, and evident injuries) is low for both cluster 1 (marginal effects −0.0228, −0.0786, and −0.415) and cluster 4 (marginal effects −0.0082, −0.0344, and −0.2374), which is probably related to the driving speed of a driver. Similar results can be found when motorists failed to yield in cluster 1 and cluster 4. In contrast, head on crashes and motorists turning or merging might increase the likelihood of suffering severe injuries for crashes occurred at non-intersection locations in daylight, and crashes caused by mid-aged drivers (25–59) and elderly drivers (60+) with low speed (20 mph), respectively, based on the marginal effects shown in Table 5. For crashes occurred at crossing paths and par-

Table 5e
Average marginal effects for cluster 5.

Variable		Cluster5				
		F	D	E	P	N
Cyclist characteristics						
Age	< 16	−0.0600	−0.0512	−0.1140	0.1612	0.0640
	16–24	−0.0476	−0.0380	−0.0697	0.1150	0.0403
	25–54	−0.0748	−0.0540	−0.0530	0.1420	0.0398
Driver characteristics						
Age	25–59	0.0294	0.0230	0.0311	−0.0656	−0.0179
Vehicle characteristics						
Veh_Type	Single Unit Truck	0.0825	0.0523	0.0112	−0.1206	−0.0254
Veh Speed	<20 mph	−0.0590	−0.0522	−0.1394	0.1754	0.0751
	20–30 mph	−0.0527	−0.0453	−0.0984	0.1447	0.0516
	30–40 mph	−0.0379	−0.0306	−0.0477	0.0896	0.0266
Crash characteristics						
Speeding	Yes	0.2066	0.0162	−0.0606	−0.2360	0.0738
Roadway characteristics						
No. of Lanes	<=4	−0.0721	−0.0471	−0.0181	0.1125	0.0247
Environmental Characteristics						
Light Condition	Daylight	−0.0298	−0.0224	−0.0234	0.0604	0.0153
	Dark - lighted roadway	−0.0498	−0.0422	−0.0968	0.1372	0.0517

N - No Injury.
P - Possible Injury.
E - Evident Injury.
D - Disabling Injury.
F - Fatal Injury.

Table 5f
Average marginal effects for cluster 6.

Variable		Cluster6				
		F	D	E	P	N
Vehicle characteristics						
Veh Speed	30–40 mph	0.0193	0.0326	0.0500	−0.0704	−0.0316
	40–50 mph	0.0340	0.0561	0.0784	−0.1201	−0.0483
	50–60 mph	0.0556	0.0837	0.0728	−0.1574	−0.0547
Crash characteristics						
Crash Type	Bicyclist Turn/Merge	−0.0162	−0.0301	−0.0644	0.0738	0.0369
	Head-On	0.0814	0.1146	0.0508	−0.1892	−0.0577
Roadway characteristics						
Road Type	Two-way	−0.3554	−0.2037	0.1903	0.2946	0.0742
Traffic Control	Yes	0.0125	0.0220	0.0386	−0.0508	−0.0223
No. of Lanes	<=4	−0.0262	−0.0428	−0.0516	0.0876	0.0330
Environmental Characteristics						
Weather	Fog, smog, smoke, rain, snow	−0.0129	−0.0239	−0.0532	0.0591	0.0310

N - No Injury.
P - Possible Injury.
E - Evident Injury.
D - Disabling Injury.
F - Fatal Injury.

allel paths, bicyclists are more likely to have possible injury or no injury.

Speeding can affect bicyclist injury severity differently in cluster 3–5. For crashes that occurred on not lighted non-intersection locations in dark (cluster 3), and crashes caused by pickups at non-intersections (cluster 5), speeding could increase the probability of fatal and disabling injuries, while for crashes that occurred on lighted roadways in dark (cluster 4), speeding could only be more likely to cause possible injury or even no injury. Comparing cluster 3 and cluster 4, it can be concluded that lighting condition could be a critical impact factor to the severe injury severities. Therefore, it is important to provide better light condition to reduce the likelihood of bicyclists suffering from severe injuries. Street lights are recommended to be built to ensure good light condition for bicyclists.

For crashes that occurred in urban areas, bicyclists are less likely to suffer severe injuries including fatal, disabling, and evident injuries (marginal effects −0.0006, −0.0111, and −0.0954) in cluster 2 (i.e., crashes caused by mid-aged drivers with low vehicle speed). This is reasonable since low speed usually has negative impacts on severe injuries. In addition, roadways in urban areas have better access control and road conditions.

5.2.5. Roadway characteristics

Curved roadway has a positive impact on severe injury severities, especially for cluster 1, cluster 2, and cluster 7, which indicates a higher probability of severe injuries for particular crash patterns including crashes occurred with vehicle speed from 20 to 30 mph when the bicyclists failed to yield and crashes caused by mid-aged drivers (25+) with low vehicle speed (<20 mph).

Table 5g
Average marginal effects for cluster 7.

Variable		Cluster7				
		F	D	E	P	N
Cyclist characteristics						
Alcohol usage	Yes	0.0366	0.0552	0.3050	−0.3259	−0.0708
Crash characteristics						
Bike Direction	Facing traffic	−0.0021	−0.0036	−0.0749	0.0542	0.0264
Crash Type	Motorist Turn/Merge	0.0038	0.0065	0.1159	−0.0911	−0.0350
	Crossing Paths	−0.0050	−0.0087	−0.2565	0.1117	0.1585
	Parallel Paths	−0.0059	−0.0104	−0.3518	−0.0408	0.4089
Roadway characteristics						
Road Geometry	Curve	0.0082	0.0136	0.1720	−0.1482	−0.0456
Divided Road	Yes	−0.0026	−0.0044	−0.0942	0.0651	0.0360
Temporal Characteristics						
Crash Time	10:00–14:59	−0.0039	−0.0066	−0.1313	0.0937	0.0480
	15:00–17:59	−0.0036	−0.0061	−0.1191	0.0828	0.0460

N - No Injury.
P - Possible Injury.
E - Evident Injury.
D - Disabling Injury.
F - Fatal Injury.

Table 5h
Average marginal effects for the whole dataset.

Variable		Whole Data				
		F	D	E	P	N
Cyclist characteristics						
Age	<16	−0.0265	−0.0098	−0.0162	0.0246	0.0279
	16–24	−0.0193	0.0024	−0.0290	0.0173	0.0286
	25–54	−0.0065	−0.0112	−0.0407	0.0381	0.0202
Alcohol usage	Yes	0.0175	0.0256	0.0069	−0.0634	0.0134
Driver characteristics						
Age	<25	0.0066	0.0118	0.0394	−0.0390	−0.0188
	25–59	0.0033	0.0059	0.0216	−0.0201	−0.0107
Alcohol usage	Yes	0.0714	0.0464	−0.0184	−0.0590	−0.0404
Vehicle characteristics						
Veh Speed	<20 mph	−0.0365	−0.0606	−0.0276	0.0964	0.0282
	20–30 mph	−0.0197	−0.0278	0.0345	−0.0009	0.0140
	40–50 mph	0.0104	0.0189	0.0609	−0.0629	−0.0273
	50–60 mph	0.0390	0.0461	0.0223	−0.0934	−0.0140
Veh_Type	Pickup	0.0080	0.0112	0.0071	0.0191	−0.0454
	Van	0.0067	0.0117	0.0379	−0.0385	−0.0178
	Single Unit Truck	0.0234	0.0390	0.0961	−0.1155	−0.0430
Crash characteristics						
Bike Direction	Facing traffic	−0.0042	−0.0245	−0.0834	0.0986	0.0135
Speeding	Yes	0.0089	0.0156	0.0483	−0.0505	−0.0223
Crash Type	Bicyclist Failed to Yield	0.0071	0.0652	−0.0020	−0.0555	−0.0148
	Head-On	0.0240	0.0402	0.0992	−0.1190	−0.0443
	Motorist Turn/Merge	−0.0137	0.0020	0.1119	−0.0721	−0.0280
Roadway characteristics						
Road Geometry	Curve	0.0124	0.0215	0.0643	−0.0690	−0.0292
Road Type	Two-way	0.0065	0.0121	0.0492	−0.0420	−0.0258
Temporal Characteristics						
Crash Time	10:00–14:59	−0.0032	−0.0059	−0.0218	0.0201	0.0108

N - No Injury.
P - Possible Injury.
E - Evident Injury.
D - Disabling Injury.
F - Fatal Injury.

Different effects are identified for the impact of two-way road on bicyclist injury severities. To be specific, this factor could increase the likelihood of severe injuries including fatal, disabling, and evident injuries in cluster 1, while decreasing the probability of fatal and disabling injuries in cluster 6. This result indicates the necessity of conducting latent class clustering analysis to reveal and interpret the variation of the effect of variables across different clusters.

Furthermore, divided roadway is found to be less likely to result in severe injuries (fatal, disabling, and evident injuries) in cluster 7 with marginal effect being −0.0026, −0.0044, and −0.0942, respectively. Therefore, in order to improve cycling safety, constructing divided roadways is recommended. Bad road conditions could result in different bicyclist injury outcomes in cluster 2 and cluster 3. Bicyclists might be more likely to be severely injured in cluster 2, while being less likely to have severe injuries in cluster 3, which indicates the heterogeneity between clusters. Some similarities can be found in cluster 2 and cluster 6, where traffic control has a positive impact on the high level of bicyclist injury severity. This is probably because traffic control is always related to the intersection areas, which may be a reason for severe injuries outcomes. For the impact of the number of lanes on bicyclist injury severity, a smaller number of lanes has a negative effect on severe injuries, which means that severe bicyclist injuries could be less likely to occur on fewer lanes, especially in cluster 5 and cluster 6.

5.2.6. Temporal characteristics

Different times of day could have various impacts on bicyclist injury severity according to the marginal effects in Table 5. Biking during midnight (0:00–5:59) may increase the likelihood of suffering severe injury severities including fatal, disabling, and evident injuries especially in cluster 2. Also, in cluster 2, different effects can be found during 6:00 to 9:59 in the morning, when bicyclists are less likely to be severely injured. Similar trend can be seen in cluster 7, where bicyclists are at a lower risk to get severe injuries during midday from 10:00 to 14:59. However, for crashes occurred during 15:00 to 17:59, bicyclists are more likely to be fatally injured in cluster 3, while being less likely to have fatal injury in cluster 7, which shows the heterogeneity features between different clusters. This result might correspond to the characterization of cluster 7, which is described as crashes caused by elderly drivers (60+) with low vehicle speed (<20 mph). Summarizing the results regarding the various impact of temporal characteristics on bicyclist injury severity for different clusters, several findings can be concluded. First, temporal characteristics are to a great extent associated with the light condition. The effects of the time periods reflect a part of the influence of light condition in another aspect. Second, the traffic volume varies during different time periods, which is related to the bicyclist injury severities.

5.2.7. Environmental characteristics

Adverse weather conditions are found to have a negative impact on bicyclist injury severity in cluster 2 and cluster 6, while it could increase the probability of severe injuries in cluster 3 (marginal effects 0.089, 0.663, and 0.0247). Since cluster 3 is characterized as crashes occurring on not lighted non-intersection locations in dark, the high probability of severe bicyclist injuries might result from the dark lighting condition. It can be inferred that adverse weather conditions are not the determining factors to severe injuries.

Lighting condition is a significant impact factor to bicyclist injury severity especially for cluster 5. Compared to dark, not lighted roadways, daylight and lighted roadways could decrease the likelihood of severe bicyclist injuries in cluster 5. Maintaining clear sight is essential to enhance cycling safety. As previously mentioned, building appropriate street lights and wearing reflective materials may decrease the risk of collisions.

6. Conclusions and recommendations

This study aims to investigate the differences of the effects of impact factors on bicyclist injury severity existing in various crash patterns. Since bicyclists are more vulnerable compared to other road users, it is essential to evaluate the variables contributing to severe bicyclist injuries. Based on the police reported data collected from 2007 to 2014 in North Carolina, latent class clustering analysis and the subsequent partial proportional odds models and/or ordered logit models are developed. Seven clusters are identified, and four PPO models and three ORL models are built to observe and interpret the impacts on bicyclist injury severity for certain patterns of crashes. It is tested that the sub-models have better goodness of fit compared to the single model developed with the whole dataset, which is consistent with the results of other research studies conducted by Depaire et al. (2009) and Sun et al. (2019).

To better analyze and interpret the bicyclist injury outcomes, marginal effects are calculated to reveal the accurate effect of each significant variable. Results show that some variables only have significant influence in specific clusters, and the impact of the same variable for various clusters can be different, which indicate that latent class clustering provides as a more accurate and insightful method to explore the information on the impact of contributing factors for further analysis of bicyclist injury severity and the improvement of bicyclist safety.

In addition, the findings of the model results provide some more useful information and guidance to help mitigate severe bicyclist injuries and improve biking safety. To be specific, since bicyclists under the influence of alcohol are found to be more likely to suffer severe injuries, it is critical to establish relative regulations to prevent bicyclists from drinking alcohol while riding a bicycle. Similar enforcement should be enhanced for drivers to inhibit driving under the influence of alcohol. Furthermore, the lighting condition is a significant contributing factor to severe injuries. Therefore, constructing lights in the areas with high bicyclist activities might help improve biking safety.

However, there are still some limitations in this research study. It should be noted that it is possible to model the latent class and the bicyclist injury severity simultaneously, for example, by combining latent class and multinomial logit models that could be more effectively compared to the sequential approach adopted in this study (Yasmin et al., 2014a, 2014b). In addition, although latent class clustering tries to maximize the homogeneity within clusters, the unobserved heterogeneity might be neglected when developing PPO/ORL models. Therefore, in the future study, the authors will try to apply mixed effect PPO model (Eluru & Yasmin, 2015) for the analysis of bicyclist injury severity.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors want to express their deepest gratitude for the financial support by the United States Department of Transportation, University Transportation Center through the Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE) at The University of North Carolina at Charlotte (Grant Number: 69A3551747133).

References

Behnood, A., & Mannering, F. (2017). Determinants of bicyclist injury severities in bicycle-vehicle crashes: A random parameters approach with heterogeneity in means and variances. *Analytic Methods in Accident Research, 16*, 35–47.
Biernacki, C., & Govaert, G. (1999). Choosing models in model-based clustering and discriminant analysis. *Journal of Statistical Computation and Simulation, 64*(1), 49–71.
Bijmolt, T. H., Paas, L. J., & Vermunt, J. K. (2004). Country and consumer segmentation: Multi-level latent class analysis of financial product ownership. *International Journal of Research in Marketing, 21*, 323–340.

Chang, F., Xu, P., Zhou, H., Chan, A. H., & Huang, H. (2019). Investigating injury severities of motorcycle riders: A two-step method integrating latent class cluster analysis and random parameters logit model. *Accident Analysis & Prevention, 131*, 316–326.

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. John Wiley & Sons.

Decker, S., Otte, D., Cruz, D. L., Müller, C. W., Omar, M., Krettek, C., & Brand, S. (2016). Injury severity of pedestrians, bicyclists and motorcyclists resulting from crashes with reversing cars. *Accident Analysis & Prevention, 94*, 46–51.

Depaire, B., Wets, G., & Vanhoof, K. (2008). Traffic accident segmentation by means of latent class clustering. *Accident Analysis & Prevention, 40*(4), 1257–1266.

Eluru, N., & Yasmin, S. (2015). A note on generalized ordered outcome models. *Analytic Methods in Accident Research, 8*, 1–6.

Eluru, N. (2013). Evaluating alternate discrete choice frameworks for modeling ordinal discrete variables. *Accident Analysis & Prevention, 55*, 1–11.

Eluru, N., Bhat, C. R., & Hensher, D. A. (2008). A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis & Prevention, 40*(3), 1033–1054.

Gårder, P. (1994). Bicycle accidents in Maine: An analysis. *Transportation Research Record*. 34-34.

Gårder, P., Leden, L., & Pulkkinen, U. (1998). Measuring the safety effect of raised bicycle crossings using a new research methodology. *Transportation Research Record, 1636*(1), 64–70.

Gong, L., & Fan, W. D. (2017). Modeling single-vehicle run-off-road crash severity in rural areas: Accounting for unobserved heterogeneity and age difference. *Accident Analysis & Prevention, 101*, 124–134.

Götschi, T., Garrard, J., & Giles-Corti, B. (2016). Cycling as a part of daily life: A review of health perspectives. *Transport Reviews, 36*(1), 45–71.

Kaplan, S., Vavatsoulas, K., & Prato, C. G. (2014). Aggravating and mitigating factors associated with cyclist injury severity in Denmark. *Journal of Safety Research, 50*, 75–82.

Kashani, A. T., & Mohaymany, A. S. (2011). Analysis of the traffic injury severity on two-lane, two-way rural roads based on classification tree models. *Safety Science, 49*(10), 1314–1320.

Kelly, P., Kahlmeier, S., Götschi, T., Orsini, N., Richards, J., Roberts, N., ... Foster, C. (2014). Systematic review and meta-analysis of reduction in all-cause mortality from walking and cycling and shape of dose response relationship. *International Journal of Behavioral Nutrition and Physical Activity, 11*(1), 132.

Kim, J. K., Kim, S., Ulfarsson, G. F., & Porrello, L. A. (2007). Bicyclist injury severities in bicycle-motor vehicle accidents. *Accident Analysis & Prevention, 39*(2), 238–251.

Klassen, J., El-Basyouny, K., & Islam, M. T. (2014). Analyzing the severity of bicycle-motor vehicle collision using spatial mixed logit models: A City of Edmonton case study. *Safety Science, 62*, 295–304.

Kröyer, H. R. (2015). The relation between speed environment, age and injury outcome for bicyclists struck by a motorized vehicle–a comparison with pedestrians. *Accident Analysis & Prevention, 76*, 57–63.

Lanza, S. T., & Rhoades, B. L. (2013). Latent class analysis: An alternative perspective on subgroup analysis in prevention and treatment. *Prevention Science, 14*(2), 157–168.

Lanza, S. T., Collins, L. M., Lemmon, D. R., & Schafer, J. L. (2007). PROC LCA: A SAS procedure for latent class analysis. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(4), 671–694.

Li, Y., & Fan, W. (2019). Pedestrian injury severities in pedestrian-vehicle crashes and the partial proportional odds logit model: Accounting for age difference. *Transportation Research Record, 2673*(5), 731–746.

Lin, Z., & Fan, W. (2019a). Cyclist injury severity analysis with mixed-logit models at intersections and nonintersection locations. *Journal of Transportation Safety & Security*, 1–23.

Lin, Z., & Fan, W. (2019b). Modeling bicyclist injury severity in bicycle–motor vehicle crashes that occurred in urban and rural areas: A mixed logit analysis. *Canadian Journal of Civil Engineering, 46*(10), 924–933.

Liu, J., Khattak, A. J., Li, X., Nie, Q., & Ling, Z. (2020). Bicyclist injury severity in traffic crashes: A spatial approach for geo-referenced crash data to uncover non-stationary correlates. *Journal of Safety Research*.

Liu, P., & Fan, W. (2018). Modeling head-on crash severity on NCDOT freeways: A mixed logit model approach. *Canadian Journal of Civil Engineering, 46*(6), 322–328.

Liu, P., & Fan, W. (2020). Exploring injury severity in head-on crashes using latent class clustering analysis and mixed logit model: A case study of North Carolina. *Accident Analysis & Prevention, 135* 105388.

Macmillan, A., Connor, J., Witten, K., Kearns, R., Rees, D., & Woodward, A. (2014). The societal costs and benefits of commuter bicycling: Simulating the effects of specific policies using system dynamics modeling. *Environmental Health Perspectives, 122*(4), 335–344.

Mannering, F. L., & Bhat, C. R. (2014). Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research, 1*, 1–22.

McLachlan, G. J., & Peel, D. (2000). *Finite mixture models*. New York: Wiley.

Mohamed, M. G., Saunier, N., Miranda-Moreno, L. F., & Ukkusuri, S. V. (2013). A clustering regression approach: A comprehensive injury severity analysis of pedestrian–vehicle crashes in New York, US and Montreal, Canada. *Safety Science, 54*, 27–37.

Mooradian, J., Ivan, J. N., Ravishanker, N., & Hu, S. (2013). Analysis of driver and passenger crash injury severity using partial proportional odds models. *Accident Analysis & Prevention, 58*, 53–58.

Moore, D. N., Schneider Iv, W. H., Savolainen, P. T., & Farzaneh, M. (2011). Mixed logit analysis of bicyclist injury severity resulting from motor vehicle crashes at intersection and nonintersection locations. *Accident Analysis & Prevention, 43*(3), 621–630.

Nabors, D., Goughnour, E., Thomas, L., DeSantis, W., & Sawyer, M. (2012). *Bicycle road safety audit guidelines and prompt lists (No. FHWA-SA-12-018)*. United States: Federal Highway Administration. Office of Safety.

NHTSA (2019). *Traffic Safety Facts 2017: A compilation of motor vehicle crash data from the fatality analysis reporting system and the General Estimates System*. Washington DC: U. S Department of Transportation.

Nilsson, P., Stigson, H., Ohlin, M., & Strandroth, J. (2017). Modelling the effect on injuries and fatalities when changing mode of transport from car to bicycle. *Accident Analysis & Prevention, 100*, 30–36.

Pai, C. W., & Saleh, W. (2007). An analysis of motorcyclist injury severity under various traffic control measures at three-legged junctions in the UK. *Safety Science, 45*(8), 832–847.

Peterson, B., & Harrell, F. (1990). Partial proportional odds models for ordinal response variables. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 39*(2), 205–217.

Prati, G., Pietrantoni, L., & Fraboni, F. (2017). Using data mining techniques to predict the severity of bicycle crashes. *Accident Analysis & Prevention, 101*, 44–54.

Pucher, J., & Dijkstra, L. (2003). Promoting safe walking and cycling to improve public health: Lessons from the Netherlands and Germany. *American Journal of Public Health, 93*(9), 1509–1516.

Pucher, J., Buehler, R., & Seinen, M. (2011). Bicycling renaissance in North America? An update and re-appraisal of cycling trends and policies. *Transportation Research Part A: Policy and Practice, 45*(6), 451–475.

Rash-ha Wahi, R., Haworth, N., Debnath, A. K., & King, M. (2018). Influence of type of traffic control on injury severity in bicycle–motor vehicle crashes at intersections. *Transportation Research Record, 2672*(38), 199–209.

Robartes, E., & Chen, T. D. (2017). The effect of crash characteristics on cyclist injuries: An analysis of Virginia automobile-bicycle crash data. *Accident Analysis & Prevention, 104*, 165–173.

Rojas-Rueda, D., de Nazelle, A., Tainio, M., & Nieuwenhuijsen, M. J. (2011). The health risks and benefits of cycling in urban environments compared with car use: Health impact assessment study. *BMJ, 343* d4521.

Sasidharan, L., & Menéndez, M. (2014). Partial proportional odds model - An alternate choice for analyzing pedestrian crash injury severities. *Accident Analysis & Prevention, 72*, 330–340.

Sasidharan, L., Wu, K.-F., & Menendez, M. (2015). Exploring the application of latent class cluster analysis for investigating pedestrian crash injury severities in Switzerland. *Accident Analysis & Prevention, 85*, 219–228.

Shaheed, M. S., & Gkritza, K. (2014). A latent class analysis of single-vehicle motorcycle crash severity outcomes. *Analytic Methods in Accident Research, 2*, 30–38.

Sivasankaran, S. K., & Balasubramanian, V. (2020). Exploring the severity of bicycle–Vehicle crashes using latent class clustering approach in India. *Journal of Safety Research, 72*, 127–138.

Stone, M., & Broughton, J. (2003). Getting off your bike: Cycling accidents in Great Britain in 1990–1999. *Accident Analysis & Prevention, 35*(4), 549–556.

Sun, M., Sun, X., & Shan, D. (2019). Pedestrian crash analysis with latent class clustering method. *Accident Analysis & Prevention, 124*, 50–57.

Ulfarsson, G. F., & Mannering, F. L. (2004). Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. *Accident Analysis & Prevention, 36*(2), 135–147.

Valent, F., Schiava, F., Savonitto, C., Gallo, T., Brusaferro, S., & Barbone, F. (2002). Risk factors for fatal road traffic accidents in Udine, Italy. *Accident Analysis & Prevention, 34*(1), 71–84.

Vanparijs, J., Panis, L. I., Meeusen, R., & de Geus, B. (2015). Exposure measurement in bicycle safety analysis: A review of the literature. *Accident Analysis & Prevention, 84*, 9–19.

Wang, X., & Abdel-Aty, M. (2008). Analysis of left-turn crash injury severity by conflicting pattern using partial proportional odds models. *Accident Analysis & Prevention, 40*(5), 1674–1682.

Washington, S. P., Karlaftis, M. G., & Mannering, F. (2010). *Statistical and econometric methods for transportation data analysis*. Chapman and Hall/CRC.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT Press.

Xia, T., Zhang, Y., Crabb, S., & Shah, P. (2013). Cobenefits of replacing car trips with alternative transportation: a review of evidence and methodological issues. Journal of Environmental and Public Health, 2013.

Yan, X., Ma, M., Huang, H., Abdel-Aty, M., & Wu, C. (2011). Motor vehicle-bicycle crashes in Beijing: Irregular maneuvers, crash patterns, and injury severity. *Accident Analysis & Prevention, 43*(5), 1751–1758.

Yasmin, S., Eluru, N., & Ukkusuri, S. V. (2014a). Alternative ordered response frameworks for examining pedestrian injury severity in New York City. *Journal of Transportation Safety & Security, 6*(4), 275–300.

Yasmin, S., Eluru, N., Bhat, C. R., & Tay, R. (2014b). A latent segmentation based generalized ordered logit model to examine factors influencing driver injury severity. *Analytic Methods in Accident Research, 1*, 23–38.

Yu, H., Li, Z., Zhang, G., & Liu, P. (2019). A latent class approach for driver injury severity analysis in highway single vehicle crash considering unobserved heterogeneity and temporal influence. *Analytic Methods in Accident Research, 24* 100110.

Zijing Lin is a Ph.D. student in the Infrastructure and Environmental Systems (INES) program at the University of North Carolina at Charlotte, Charlotte, NC, U.S.A. She is with the Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE). Her research topics include transportation safety data analysis, and non-motorized transportation.

Wei (David) Fan currently serves as an associate professor in the Department of Civil and Environmental Engineering at The University of North Carolina at Charlotte (UNC Charlotte). He is the Director of the USDOT Tier I University Transportation Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE). Dr. Fan holds a Ph.D. (May 2004) in Civil Engineering – Transportation from the University of Texas at Austin (Hook 'em Horns!). He was a Senior Analytical Optimization Software Developer for the R&D Department at SAS Institute Inc. located in Cary, North Carolina from June 2004 – August 2006. Dr. Fan's primary research interests include operations research, transportation network modeling & optimization, multimodal transportation systems planning and operations, traffic operations and simulation, transportation safety and crash analysis, intelligent transportation systems (ITS), and computer software development. Dr. Fan serves as an associate editor of IEEE Transactions – Intelligent Transportation Systems while also serving as a member of three other transportation journal editorial boards. Dr. Fan also serves as a member on the National Science Foundation (NSF) review panels, two National Cooperative Highway Research Program (NCHRP) panels, one Transit Cooperative Research Program (TCRP) panel, and three American Society of Civil Engineers (ASCE) Committees (on the Connected & Autonomous Vehicles Impacts, Public Transport, and Rail Transportation), as well as two Transportation Research Board (TRB) committees. He has been working as an active ad hoc reviewer for more than 20 prestigious journals as well as the national Highway Capacity Manual (HCM) (2010 version). He was selected by the students as the Best Civil Engineering Professor of the Year in 2007 at The University of Texas at Tyler and was a proud ASCE Excellence in Civil Engineering Education (ExCEEd) Teaching Fellow in 2007.

Dr. Fan is involved in many sponsored projects (over 14 million dollars in funding), having been principal or co-principal investigator on many research studies for the U. S. Department of Transportation (USDOT), Federal Highway Administration (FHWA), National Cooperative Highway Research Program (NCHRP), SHRP2 Education Connection, Texas Department of Transportation (TxDOT) and North Carolina Department of Transportation (NCDOT). He had published more than 100 journal articles, proceeding papers and technical reports on transportation network modeling, multimodal transportation systems planning and operations, traffic operations and simulations, traffic safety and crash analysis, and operations research applications in transportation. He is a registered professional engineer in Texas.

Journal of Safety Research 76 (2021) 184–196

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Exploring pedestrian injury severities at pedestrian-vehicle crash hotspots with an annual upward trend: A spatiotemporal analysis with latent class random parameter approach

Li Song [a], Wei (David) Fan [a,*], Yang Li [a], Peijie Wu [b]

[a] USDOT Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE), Department of Civil and Environmental Engineering, University of North Carolina at Charlotte, EPIC Building, Room 3366, 9201 University City Boulevard, Charlotte, NC 28223-0001, United States
[b] School of Transportation Science and Engineering, Harbin Institute of Technology, No. 73, Huanghe Road, Nangang District, Harbin, China

A R T I C L E   I N F O

Article history:
Received 10 August 2020
Received in revised form 1 September 2020
Accepted 9 December 2020
Available online 29 December 2020

Keywords:
Pedestrian
Injury severity
Spatiotemporal analysis
Latent class clustering
Random parameter logit model

A B S T R A C T

Introduction: With the increasing trend of pedestrian deaths among all traffic fatalities in the past decade, there is an urgent need for identifying and investigating hotspots of pedestrian-vehicle crashes with an upward trend. *Method:* To identify pedestrian-vehicle crash locations with aggregated spatial pattern and upward temporal pattern (i.e., hotspots with an upward trend), this paper first uses the average nearest neighbor and the spatial autocorrelation tests to determine the grid distance and the neighborhood distance for hotspots, respectively. Then, the spatiotemporal analyses with the Getis-Ord Gi* index and the Mann-Kendall trend test are utilized to identify the pedestrian-vehicle crash hotspots with an annual upward trend in North Carolina from 2007 to 2018. Considering the unobserved heterogeneity of the crash data, a latent class model with random parameters within class is proposed to identify specific contributing factors for each class and explore the heterogeneity within classes. Significant factors of the pedestrian, vehicle, crash type, locality, roadway, environment, time, and traffic control characteristics are detected and analyzed based on the marginal effects. *Results:* The heterogeneous results between classes and the random parameter variables detected within classes further indicate the superiority of latent class random parameter model. *Practical Applications:* This paper provides a framework for researchers and engineers to identify crash hotspots considering spatiotemporal patterns and contribution factors to crashes considering unobserved heterogeneity. Also, the result provides specific guidance to developing countermeasures for mitigating pedestrian-injury at pedestrian-vehicle crash hotspots with an upward trend.

1. Introduction

Compared to other entities in traffic crashes, pedestrians are more vulnerable to suffer severe injuries. According to one report from the National Highway Safety Administration (NHTSA, 2019), in the United States there were 5,977 pedestrian fatalities in traffic crashes in 2017. From 2008 to 2017, the percentage of pedestrian deaths in total traffic fatalities has constantly increased from 12% to 16%. In recent years, more and more efforts have been put into investigating contributing factors of the pedestrian injury severity at specific hazardous locations (Anderson, 2009; Dai, 2012). Meanwhile, existence of the temporal variation and tendency of the pedestrian crash data might affect the model result in different ways (Behnood & Mannering, 2016), and neglecting the fundamental temporal features could result in erroneous conclusions (Mannering, 2018). One previous research study has identified the instability of different time scales among the pedestrian crashes, and the annual variation mainly shows an increasing/decreasing trend (Dai, 2012). Hence, there is an urgent need to develop a proper approach to identifying the contributing factors at crash hotspots with annual uptrends.

Previous studies have applied several methods to explore the factors to crash severity. A detailed review was summarized in (Mannering & Bhat, 2014), and this review also pointed out that the heterogeneity inherent in the crash observations could result in biased parameter estimations and incorrect inferences. To obtain more accurate and specific model results, it is important to investigate the pedestrian injury severity by considering the heterogeneity both within and between the pedestrian crash observations.

* Corresponding author.
E-mail address: wfan7@uncc.edu (Wei (David) Fan).

https://doi.org/10.1016/j.jsr.2020.12.008
0022-4375/

To identify specific contributing factors and provide guidance for improving the deteriorative tendency of pedestrian-vehicle crashes at hotspots, this paper uses the spatiotemporal trend analysis with the Getis-Ord Gi* index and the Mann-Kendall trend test to explore the annual spatial clustering and the temporal tendency of pedestrian-vehicle crashes in North Carolina from 2007 to 2018. Meanwhile, the grid distance interval and the neighborhood distance are determined by the average nearest neighbor and the spatial autocorrelation test, respectively. Then, a sequential process by combining latent class clustering with random parameter logit approaches are used to identify contributing factors considering the heterogeneity within and between the classes.

2. Literature review

2.1. Crash locations considering spatiotemporal patterns

To identify aggregated/high-frequency traffic crash locations, point pattern analyses, such as the kernel density estimation (KDE) (Ouni & Belloumi, 2018) and Getis-Ord Gi* index (Songchitruksa & Zeng, 2010), were commonly used in previous studies. However, KDE is not feasible for the statistical significance test and the density pattern will certainly be influenced by the choice of bandwidth (Plug, Xia, & Caulfield, 2011). Hence, the Getis-Ord Gi* index, which is a statistical-based test for high/low value clusters, was then deployed to identify the spatial patterns in several studies. Ulak, Kocatepe, Ozguven, Horner, and Spainhour (2017) employed the Getis-Ord Gi* index to identify hotspots with the optimized neighborhood distance that was determined by the Global Moran's I test. Results showed that the accessibility to hospitals of hotspots is one of the major reasons for severe injuries. Considering that the Getis-Ord Gi* and the global Moran's I could identify spatial patterns from local and global perspectives, respectively (Blazquez, Picarte, Calderón, & Losada, 2018), this paper employs the Getis-Ord Gi* for hotspots identification and utilizes the global Moran's I to provide a reference for the neighborhood distance. Meanwhile, the average nearest neighbor (ANN) was employed to calculate the distance interval between traffic crashes (Yalcin & Sebnem Duzgun, 2015).

For the temporal trend analysis, the Mann-Kendall trend test, which is a statistical-based non-parametric rank correlation analysis method, has been widely used in previous studies (Gudes, Varhol, Sun, & Meuleners, 2017; Wang & Chan, 2016). Gudes et al. (2017) evaluated temporal patterns of the hot/cold spot regions of the heavy-vehicle crashes by the Mann-Kendall trend test. Results showed inconsistency of temporal patterns in hotspots over time. With such analyses, the temporal tendency of hotspots could be further investigated.

2.2. Identification of injury-severity factors considering unobserved heterogeneity

As summarized in Table 1, statistics-based methods, such as an ordered/unordered response model with a logit/probit link function, have been widely used because of their good performance in parameter calibration and outcome interpretation (Mannering & Bhat, 2014). Moreover, to avoid biased parameters estimation and incorrect inferences caused by the unobserved heterogeneity, random parameter models, which can potentially capture unobserved heterogeneity by allowing parameters to vary across observations, were proposed (Mannering & Bhat, 2014). Abay (2013) compared the pedestrian severity outcomes with ordered logit, mixed ordered logit, multinomial logit, and mixed logit. The result revealed that mixed models can accommodate flexible variable effects to some extent while fixed-parameters injury severity models underestimated the effect of some important behavioral attributes of the crashes.

For a heterogeneity-based data segmentation approach employed in pedestrian injury severities, Table 1 also shows many sequential processes of combining the Latent Class Clustering (LCC) with other models, such as the Multinomial Logit model (MNL) (Sun, Sun, & Shan, 2019), Partial Proportional Odds model (PPO) (Li & Fan, 2019a), and Mixed Logit Model (Behnood & Mannering, 2016). Iranitalab and Khattak (2017) compared the crash severity prediction performance of the LCC and k-means clustering with the MNL and three machine learning methods. Results indicated that LCC could well improve the performance of the multinomial logit model. Behnood and Mannering (2016) analyzed differing injury-severity levels sustained by pedestrians in Chicago using both latent class and mixed logit models, which better accounts for unobserved heterogeneity compared to conventional models. Hence, a random parameter model (mixed logit model), which accounts for the heterogeneity across the observations, is considered after the implementation of LCC.

3. Methodology

3.1. Spatiotemporal analysis

3.1.1. Spatiotemporal trend analysis

The basic idea of conducting the spatiotemporal trend analysis is to first divide the map into square bins with a specific distance interval and time interval. The Getis-Ord Gi* index (Getis & Ord, 2010) and the Mann-Kendall test (Kendall & Gibbons, 1990; Mann, 1945) are used to investigate spatial hot/cold (i.e., aggregating of high/low values) pattern and the temporal tendency of these patterns, respectively. The formula of Getis-Ord Gi* index is:

$$G_i^* = \frac{\sum_{j=1}^{n}\omega_{i,j}x_j - \overline{X}\sum_{j=1}^{n}\omega_{i,j}}{SD(x_j)\sqrt{\frac{n}{n-1}\sum_{j=1}^{n}\omega_{i,j}^2 - \frac{1}{n-1}\left(\sum_{j=1}^{n}\omega_{i,j}\right)^2}} \tag{1}$$

where x_j represents the attribute value of the jth bin. $\omega_{i,j} = 1$ if the jth bin is within the spatiotemporal neighborhood distance of the ith bin and 0 otherwise. n denotes the total number of bins within the spatiotemporal neighborhood distance. $\overline{X}$ is the mean value for x_j. $SD(x_j)$ means the standard deviation for x_j. G_i^* is also a Z-score. When the p-value is statistical significance, $G_i^* > 0$ represents a clustering pattern of the high values (hotspot), $G_i^* = 0$ denotes a random pattern of the values, and $G_i^* < 0$ means a clustering pattern of the lower values (cold spot).

Then the Mann-Kendall trend test is performed on every location/grid with data within a specified time interval. For the Gi* value within each time interval $\{N_t : t = 1, 2, \cdots, T\}$, the trend test statistic S is:

$$S = \sum_{i=1}^{T-1}\sum_{j=i+1}^{T} a_{ij} \tag{2}$$

$$a_{ij} = sign(N_j - N_i) = \begin{cases} 1 N_i < N_j \\ 0 N_i = N_j \\ -1 N_i > N_j \end{cases} \tag{3}$$

where a_{ij} is a symbolic variable which counts the rank/trend of the Getis-Ord Gi* index.

The null hypothesis for S is zero, which means no trend in the values over time. Based on the variance of the values in the bin time series, Z statistic is used for the statistical significance test.

Table 1
Summary of methodological approaches in pedestrian injury severity studies.

Model	Specific scenario	Year	Location	Data	Literature
Multinomial logit model (MNL)	–	2005–2012	North Carolina	3,553	(Chen & Fan, 2019)
Partial proportional odds model (PPO)	pedestrian	2007–2014	North Carolina	10,875	(Li & Fan, 2019b)
Support vector machine and MNL	time of day	2010–2014	California	8,573	(Mokhtarimousavi, 2019)
Binary logistic regression and tree-based models	–	2014–2016	Changsha, China	791	(Hu, Wu, Huang, Peng, & Liu, 2020)
Classification and regression tree with random forest approach.	weather	2013	Britain	14,174	(Li, Ranjitkar, Zhao, Yi, & Rashidi, 2017)
Extracted rules from Bayesian networks	urban and suburban	2009–2011	Jordan	21,852	(Mujalli, Garach, López, & Al-Rousan, 2019)
Considering unobserved heterogeneity					
Mixed logit model	signalized and non-signalized locations	2008–2010	Florida	7,630	(Haleem et al., 2015)
Mixed logit model	–	1997–2000	North Carolina	5,808	(Kim et al., 2010)
Random-parameter (mixed) logit	–	2002–2006	New York City	4,666	(Aziz et al., 2013)
Artificial neural network and random parameter ordered response models	day of week	2010–2014	California	10,146	(Mokhtarimousavi, Anderson, Azizinamini, & Hadi, 2020)
Ordered logit, mixed ordered logit, multinomial logit, mixed logit	–	1998–2009	Denmark	4,952	(Abay, 2013)
Ordered logit model, generalized ordered logit model, and latent class ordered logit model	–	2002–2006	New York City	4,701	(Yasmin et al., 2014)
Latent class clustering and MNL	whole and each cluster	2006–2015	Louisiana	14,236	(Sun et al., 2019)
Latent class clustering and binary logit	whole and each cluster	2009–2012	Switzerland	9,659	(Sasidharan et al., 2015)
Latent class with ordered probit method, k-means with MNL	whole and each cluster	2002–2006 (NYC), 2003–2006 (M)	New York City, Montreal	5,820	(Mohamed et al., 2013)
Latent class clustering and PPO	each cluster	2007–2014	North Carolina	10,875	(Li & Fan, 2019a)
Latent-class logit and mixed logit models.	period (pre-recession, recession, and post-recession)	2005–2012	Chicago	19,895	(Behnood & Mannering, 2016)
Considering spatial and temporal patterns					
Bernoulli model and logistic regression	spatial clusters	2000–2007	Georgia	7,763	(Dai, 2012)
Kernel density estimation analysis and MNL	Spatiotemporal patterns	2001–2013	Tunisia	1,922	(Ouni & Belloumi, 2018)
Geographically and temporally weighted ordinal logistic regression	–	2007–2014	North Carolina	13,854	(Liu, Hainen, Li, Nie, & Nambisan, 2019)

$$Z_S = \begin{cases} \frac{S-1}{SD(S)}, S > 0 \\ 0, S = 0 \\ \frac{S+1}{SD(S)}, S < 0 \end{cases} \tag{4}$$

When $T \geq 10$, statistic S follows the normal distribution approximately. $SD(S)$ denotes the stand error of the S. For a given confidence level α, if $|Z_S| \geq |Z_{S,1-\alpha/2}|$, then the null hypothesis is rejected. Also, $Z_S > 0$ and $Z_S < 0$ indicate the uptrend and downtrend in bin values.

3.1.2. Average nearest neighbor

The average nearest neighbor analysis is used to provide a reasonable reference for the distance interval of the grid. If the average distance of the data is less than the average distance for a hypothetical random distribution, the distribution of the data being analyzed is considered clustered (Ebdon, 1985). The average nearest neighbor ratio (ANN) can be expressed as:

$$ANN = \frac{\overline{D_O}}{\overline{D_E}} = \frac{\sum_{i=1}^{n} d_i/n}{0.5/\sqrt{n/A}} = \frac{2\sum_{i=1}^{n} d_i}{\sqrt{nA}} \tag{5}$$

$$Z_{ANN} = \frac{\overline{D_O} - \overline{D_E}}{SD} \tag{6}$$

where $\overline{D_O}$ represents the observed average distance between each data point and its nearest neighbor. $\overline{D_E}$ means the expected average distance for the randomly distributed data points. d_i denotes the distance between data i and its nearest neighboring data. n corresponds to the total number of data, and A is the area of a minimum enclosing rectangle around all data points. If ANN is less than 1, the point pattern exhibits clustering. If ANN equals 1, there has no trend. If ANN is greater than 1, the trend is dispersion.

3.1.3. Spatial autocorrelation test

The Global Moran's I is a spatial autocorrelation test and can be used to evaluate the clustered, dispersed, or random spatial pattern in observations (Moran, 1948). This paper utilizes the I index to provide a reasonable reference for the neighborhood distance of bins used in the spatial–temporal analysis.

$$I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \times C_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \times D(x_i)} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \times (x_i - \overline{x})(x - \overline{x})}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \times \frac{1}{n}\sum_{1}^{n}(x_i - \overline{x})^2} \tag{7}$$

$$Z_I = \frac{I - E(I)}{\sqrt{D(I)}} \tag{8}$$

where x_i is the attribute value of j spatial location/grid. $w_{ij} = 1$ if the jth grid is within the spatial neighborhood distance of the ith grid and 0 otherwise. C_{ij} denotes the attribute similarity matrix. $E(I) = -1/(n-1)$. $D(I) = E\left(I^2\right) - E(I)^2$. If I is positive and close to 1, it denotes the incremental spatial autocorrelation (clustered pattern) between neighborhoods; if I is equal to 0, it means a random pattern of features; and if I is less than 0, it represents a dispersed pattern of features.

3.2. Latent class random parameters model

3.2.1. Latent class clustering

The latent class clustering (LCC) is a statistical model-based approach that can classify the dataset into homogenous subsets by maximizing the heterogeneity between classes (Lanza, Collins, Lemmon, & Schafer, 2007). It is assumed that the LCC segments the whole dataset with J discrete category variables into M classes. The probability of response Y can be calculated as:

$$P(Y_i = j) = \sum_{m=1}^{M} \gamma_m \prod_{j=1}^{J} \prod_{r_j=1}^{R_j} \rho_{k,r_j|m}^{I(y_j=r_j)} \tag{9}$$

where each observation i contains J categorical variables, Y_i denotes the response of the observation i for J category, and $Y_i = 1, 2, \ldots r_j$. γ_m is the membership probability for latent class cluster m (m = 1, 2, ..., M). $\rho_{k,r_j|m}^{I(y_j=r_j)}$ represents the item-response probability that observation i has response r_j being conditioned on latent class membership m. ρ means the correspondence between observed and unobserved classes. $I(y_j = r_j)$ denotes the indicator function that equals to 1 when $y_j = r_j$, and 0 otherwise.

To determine an appropriate number of classes, four commonly used criteria including Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC), Bayesian Information Criteria (BIC) and Entropy-based Measures (EM) are utilized (Song & Fan, 2020). Smaller values of the AIC, BIC, and CAIC indicate a better clustering result. Meanwhile, the EM indicates the information quality of the cluster and closing to 1 means a better clustering result (McLachlan & Peel, 2004).

3.2.2. Random parameter logit model

Following the LCC, random parameter logit model (or mixed logit model) is developed to further explore the unobserved heterogeneity for each segmented crash data. The utility function is defined as a linear function for individual i with severity level j:

$$U_{ij} = \beta_i X_{ij} + \xi_{ij} + \varepsilon_{ij} \tag{10}$$

where X_{ij} is a vector of independent variables, β_i denotes the corresponding parameter. ξ_{ij} represents the error component that has a general distribution correlated among severity levels and heteroscedastic for individuals, and ε_{ij} is the random term with independently and identically distributed Gumbel distribution over severity levels and individuals (McFadden & Train, 2000).

The probability of individual i for injury severity j is the integral of the condition choice probability $P_i(j|\xi_{ij})$ over the distribution of ξ_{ij}.

$$P_i(j|\xi_{ij}) = \frac{exp(\beta_i X_{ij} + \xi_{ij})}{\sum_{j=1}^{J} exp(\beta_i X_{ij} + \xi_{ij})} \tag{11}$$

Since the $P_i(j)$ does not always have a closed-form solution, 200 Halton draws are used in the simulation-based maximum likelihood method for parameter estimation. All the random parameters are assumed to be normally distributed since the parameters could be positive and negative. Marginal effects are used to illustrate the impact of the explanatory variable in the changing values of severity probability outcomes (Derr, 2013).

4. Data description

The data used in this paper are obtained from the North Carolina Department of Transportation (NCDOT), which include 33,707 pedestrian-vehicle crash observations in North Carolina from 2007 to 2018. The whole spatiotemporal analysis process uses a 5% significant level. The distance interval for the temporal trend analysis is set as 382 m, which is obtained by the average nearest neighbor test (ANN ratio: 0.286; Z-score: −250.845; P-value <0.0001). The inverse neighborhood distance is set as 8,000 m, which is determined by the empirical results of the spatial autocorrelation test and the total number of hotspots detected. As shown in Fig. 1, the Moran's Index at 8,000 m is 0.36, which denotes a clustering pattern within the neighborhood distance, and the z-score reaches 322. Meanwhile, the total number of hotspots reaches 5,810 with a change rate less than 5% after 8,000 m. A total number of 17,013 pedestrian-vehicle crashes at hotspots with upward annual trend are detected and is shown in Fig. 2.

To further model the contributing factors to pedestrian injury severity at hotspots with an upward trend, 13,303 pedestrian injury observations are filtered after selecting the pedestrian with the highest injury severity in single vehicle involved crashes and deleting observations with missing variables. The pedestrian severity is classified into three levels (i.e., fatal/incapacitating injury (F/I), non-incapacitating injury (NI), and no/possible injury (N/P)) by considering both the severity features and crash frequency. As shown in Appendix Table A1, explanatory variables are classified into the human, vehicle, crash, locality and roadway, environment and time, and traffic control categories.

5. Results and discussions

5.1. Latent class clustering results

The LCC is implemented to maximize the heterogeneity between the datasets. As shown in Fig. 3. The values of AIC, CAIC, and BIC all decrease with the increase of the class numbers, and the rate of change is less than 3% after four classes. Meanwhile, the entropy value for the 4-class model reaches a local maximum of 0.91, which is close to 1 and denotes a good segmentation of the data. Hence, this paper uses the LCC to segment the crash data into four latent classes.

All explanatory variables in Appendix Table A1 are utilized in the LCC analysis. Table 2 only shows the featured variables having a proportion that is significantly different from other latent classes, while other variables are not shown in Table 3 since the proportions of them are comparatively small and less descriptive. The combination of these featured variables is utilized as a latent variable/label to describe each class. For example, according to the feature variables in class 1, about 87.78% of the crashes happened in rural areas, 49.53% occurred in state secondary routes, 56.59% happened in dark without roadway lights, and 44.79% are set with double yellow lines, no passing zone sign. Hence, class 1 could be labeled as a condition of rural, state secondary route, dark without roadway lights, and double yellow line, no passing zone sign control. Similarly, class 2 can be specified as a circumstance of urban, public vehicular area, daylight, and without traffic control. Class 3 can be described as a scenario of urban, local street and driveway, daylight, and no traffic control. Class 4 can be defined as a situation of urban, dark with lighted roadway, and no traffic control.

5.2. Random parameter logit model results

After obtaining the LCC results, the heterogeneity within the crash is further investigated with four random parameter logit models. To obtain significant variables in each latent class, all explanatory variables are first utilized as the inputs in the random parameter logit model. The Chi-square test is applied as the selection criterion for both significant fixed variables and random parameter variables at a 5% significance level. Final variable coefficient estimation results are shown in Appendix Tables A2–A5.

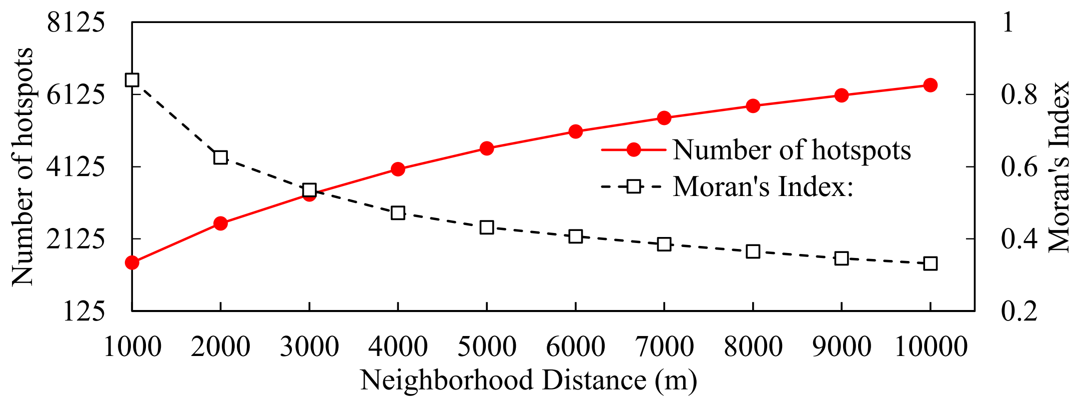


Fig. 1. Number of hotspots and the Moran's Index for different neighborhood distance.

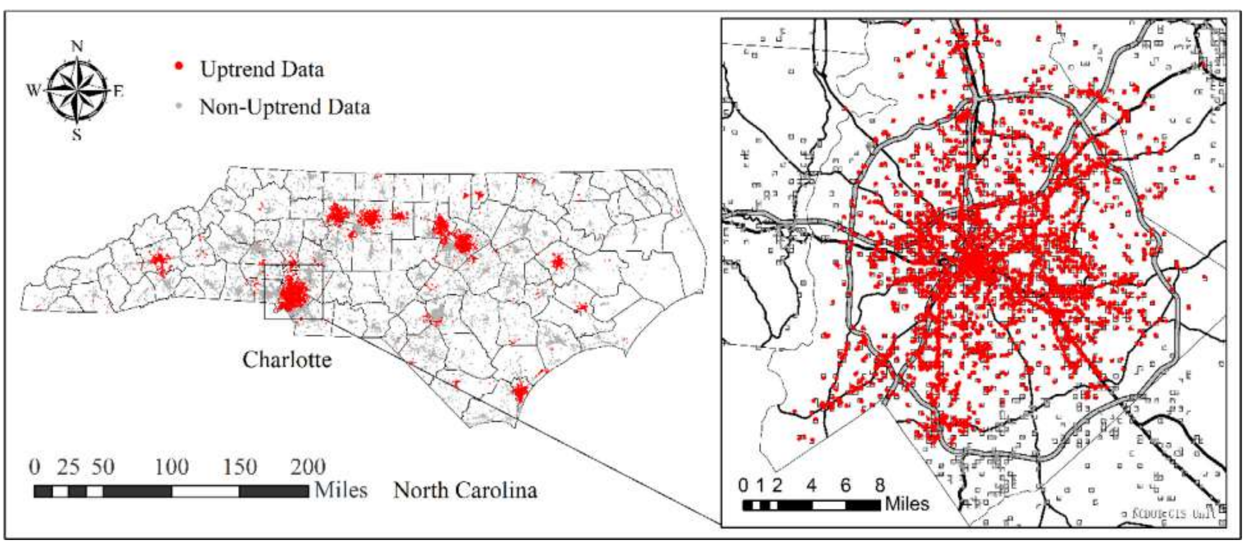


Fig. 2. Spatiotemporal trend analysis result for hotspots with upward trend in North Carolina.

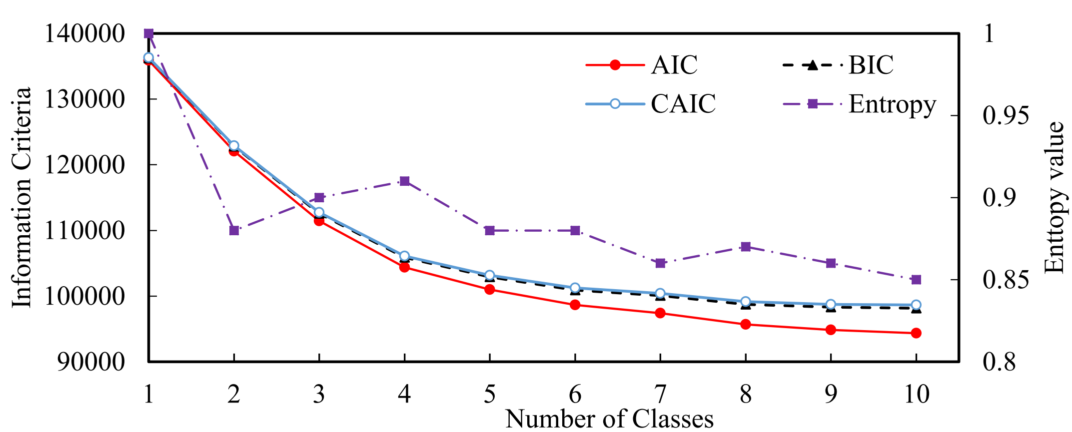


Fig. 3. Latent class results of AIC, BIC CAIC, and Entropy value for different class numbers.

Table 2
Distributions of featured variables (bold) and statistics for each latent class.

Variable	No.	Description	Class 1	Class 2	Class 3	Class 4
Locality	2	Urban	12.22%	**91.92%**	**97.64%**	**97.98%**
RdClass	4	State Secondary Route	**49.53%**	0%	0.05%	0.31%
	5	Local Street, Driveway	6.55%	9.97%	**93.95%**	**87.09%**
	6	Public Vehicular Area	1.16%	**89.89%**	1.14%	3.33%
LightCond	1	Daylight	37.05%	**77.62%**	**91.87%**	2.59%
	3	Dark – Lighted Roadway	3.06%	17.42%	2.28%	**70.86%**
	4	Dark – Roadway Not Lighted	**56.59%**	1.21%	0.22%	22.49%
Control	1	No Control Present	48.25%	**95.26%**	50.62%	66.68%
	4	Double Yellow Line, No Passing Zone	**44.79%**	0.15%	1.4%	1.7%

Table 3
Marginal effects of explanatory variables in class 1 and class 2.

Variable	Description	Class 1			Class 2		
	Severity level	F/I[a]	NI[b]	N/P[c]	F/I[a]	NI[b]	N/P[c]
PedAge2	24 < PedAge ≤ 54 (base: <24)				−0.019	−0.048	0.067
PedAge4	PedAge ≥ 65	0.231	−0.136	−0.095	−0.007	0.096	−0.090
DrvrVehTyp2	Middle (base: small)				0.022	0.029	−0.051
DrvrVehTyp3	Heavy	0.235	−0.139	−0.096	0.111	0.070	−0.181
AmbulanceR2	No ambulance rescue (base: yes)	−0.197	−0.185	0.382	−0.043	−0.143	0.186
CrashGrp2	Crossing roadway with vehicle not turning (base: walking along roadway)	0.160	−0.096	−0.064			
CrashGrp4	Off roadway				−0.010	0.142	−0.132
CrashGrp6	Dash/dart-out	0.221	−0.130	−0.090			
CrashGrp7	Backing vehicle	0.137	−0.325	0.188	−0.011	0.152	−0.142
CrashGrp10	Other/unusual circumstances				0.014	0.259	−0.272
Locality2	Urban (base: rural)				−0.025	0.009	0.016
Development2	Commercial (base: residential)				−0.015	−0.069	0.084
RdGrad2	Grade (base: level)	−0.040	0.091	−0.052			
RdClass2	Interstate (base: US route)	0.158	−0.094	−0.064			
RdClass5	Local street, driveway				0.038	0.054	−0.092
RdConfig3	Two-way, divided (base: one-way, not divided)	0.143	−0.004	−0.139			
LightCond4	Dark – roadway not lighted (base: daylight)	0.111	−0.067	−0.044			
Hour6	0:00–5:59 (base 6:00–9:59)				0.032	0.124	−0.156

Note:
[a] F/I – Fatal/ Incapacitating injury.
[b] NI – Non-incapacitating injury.
[c] N/P – No/Possible injury.

5.3. Marginal effects

The marginal effect results of the explanatory variables at a 5% significance level are shown in Tables 3 and 4. Variations of such impacts at the different severity levels and different latent classes are also detected. It is also noted that even though some variables have comparatively small proportions in the latent class subsets, they are still found to be a significant factor in the final results. The following subsections provide specific analyses and comparison for the impacts of factors on F/I and NI severity across different latent classes.

5.3.1. Pedestrian characteristics

Age and alcohol involvement are identified as significant variables. Compared to the young pedestrians (age <24), the results in classes1, 3, and 4 indicate an increasing tendency in which the probability of pedestrian suffering F/I and NI injury would both increase with the increase of the age stage. For example, the probability of pedestrians being F/I in class 4 increases from 0.035 to 0.086 and 0.175 with the increase of age stage. A similar result could be found in (Yasmin, Eluru, & Ukkusuri, 2014). However, class 2 shows heterogeneity results as the middle-age pedestrians (age within 24 to 54) and the elder pedestrians (age >65) could decrease the F/I injury by −0.019 to −0.007, respectively. Such heterogeneity within the demographical variables (age or gender) among different latent classes can also be supported by some previous studies (Abay, 2013; Aziz, Ukkusuri, & Hasan, 2013). Besides, for pedestrians under the influence of the alcohol, the possibility of pedestrians being F/I injured increases by 0.079 and 0.071 in classes 3 and 4, respectively. This result is in line with (Abay, 2013; Sasidharan, Wu, & Menendez, 2015), and more specific enforcements to limit/protect intoxicated pedestrians are needed.

5.3.2. Vehicle characteristics

This study shows that pedestrians are more vulnerable with the increase of the vehicle weight during pedestrian-vehicle crashes. For example, compared to small vehicles, the probability of pedestrians suffering F/I injury involving the middle and heavy vehicles in class 2 increases from 0.022 to 0.111, respectively.

5.3.3. Crash characteristics

Compared to ambulance rescue situations, situations without an ambulance rescue are found to have less F/I and NI injuries in all classes. This could be possibly explained by the situation under which people might not call the ambulance when the pedestrian has no/possible injury. For the hit and run situation, heterogeneous results show that there are a −0.024 probability decrease and a 0.033 probability increase for the F/I injury in classes 3 and 4,

Table 4
Marginal effects of the explanatory variables in class 3 and class 4.

Variable	Description	Class 3			Class 4		
	Severity level	F/I[a]	NI[b]	N/P[c]	F/I[a]	NI[b]	N/P[c]
PedAge2	$24 < \text{Pedage} \le 54$ (base: <24)				0.035	−0.021	−0.014
PedAge3	$55 < \text{Pedage} \le 64$	0.033	−0.015	−0.017	0.086	−0.050	−0.035
PedAge4	$\text{Pedage} \ge 65$	0.082	−0.038	−0.044	0.175	−0.102	−0.073
PedAlcFlag2	PedAlcFlag = 'yes' (base: no)	0.079	0.083	−0.161	0.071	0.024	−0.095
DrvrVehTyp2	Middle (base: small)	0.034	−0.016	−0.018			
DrvrVehTyp3	Heavy				0.160	−0.093	−0.067
HitRun2	Hit and run (base: no)	−0.024	−0.072	0.096	0.033	−0.102	0.069
AmbulanceR2	No ambulance rescue (base: yes)	−0.066	−0.248	0.315	−0.081	−0.177	0.258
CrashGrp2	Crossing roadway with vehicle not turning (base: walking along roadway)				0.089	−0.011	−0.078
CrashGrp3	Crossing roadway with vehicle turning	−0.078	−0.063	0.142	−0.105	−0.067	0.172
CrashGrp5	Pedestrian in roadway				0.109	−0.064	−0.045
CrashGrp6	Dash/dart-out	−0.030	0.204	−0.174	0.092	0.016	−0.108
CrashGrp7	Backing vehicle				0.057	−0.177	0.119
CrashGrp8	Multiple threat/trapped	−0.022	0.150	−0.128			
CrashGrp9	Bus related vehicle	−0.030	0.206	−0.176	0.212	−0.123	−0.089
Locality2	Urban (base: rural)	0.023	−0.153	0.131			
Development2	Commercial (base: residential)				0.043	−0.026	−0.018
Development4	Institutional				0.046	−0.141	0.095
RdCurve2	Curve (base: straight)	0.046	−0.021	−0.024	0.104	−0.061	−0.043
RdGrad2	Grade (base: level)	0.023	−0.011	−0.012	0.076	−0.045	−0.031
RdClass2	Interstate (base: US route)	0.046	−0.272	0.226			
RdClass5	Local street, driveway	−0.031	−0.130	0.161	−0.106	0.063	0.044
RdClass6	Public vehicular area	−0.068	−0.249	0.317	−0.089	−0.100	0.189
RdConfig2	Two-way, not divided (base: one-way, not divided)				0.024	−0.074	0.050
RdConfig3	Two-way, divided	0.020	−0.010	−0.011	0.098	−0.057	−0.041
LightCond3	Dark – lighted roadway (base: daylight)	0.197	−0.091	−0.106			
LightCond4	Dark – roadway not lighted				0.061	−0.036	−0.025
Weather2	Cloudy (base: clear)				0.073	−0.043	−0.030
Hour2	10:00–14:59 (base 6:00–9:59)	−0.029	0.014	0.015			
Hour3	15:00–17:59	−0.018	0.009	0.010	−0.096	0.058	0.038
Hour5	21:00–23:59				−0.017	0.052	−0.035
Hour6	0:00–5:59				0.053	0.039	−0.092
TraffCntrl2	Signs (base: no control)	−0.016	−0.074	0.091	0.028	−0.086	0.059
TraffCntrl4	Double yellow line, no passing zone	0.084	−0.039	−0.045	0.093	−0.055	−0.039

Note:
[a] F/I – Fatal/ Incapacitating injury.
[b] NI – Non-incapacitating injury.
[c] N/P – No/Possible injury.

respectively. Meanwhile, both two classes show a probability decrease for NI injury and a probability increase for N/P injury.

For crash type factors, the situation when the pedestrian is walking along the roadway is set as the base. Pedestrians crossing the roadway with the vehicle not turning would result in a 0.16 and 0.089 probability increase for the F/I injury in classes 1 and 4, respectively. In comparison, results show a probability decrease of −0.078 and −0.105 for the F/I injury when crossing a roadway with the turning vehicle in classes 3 and 4, respectively. One possible reason for explaining such a difference might be that the speed of the vehicles is much lower when turning than when vehicles are traveling straight. Also, the situation when vehicles are off the roadway decreases the probability with −0.01 for F/I injury in class 2. This result is also in line with Kim, Ulfarsson, Shankar, and Mannering (2010) as the driver would reduce the speed when driving off the roadway. For the situation when the pedestrian is in the roadway, a 0.109 probability increase for the F/I injury is found in class 4. A similar conclusion was also drawn in (Mohamed, Saunier, Miranda-Moreno, & Ukkusuri, 2013). In multiple threat/trapped situation, a −0.022 probability decrease for the F/I injury and a 0.15 probability increase for the NI injury are identified.

Heterogeneities also exist in variables of dash/dash-out, backing vehicle, and bus-related cases across different latent classes. In the dash/dart-out case, heterogeneity results are found in the probability of F/I injury. While, the probability of N/P injury decreases in all classes, which indicate that a severe injury outcome would occur under the dash/dart-out situation, and the results are in accord with Sun et al. (2019). Also, for the backing vehicle case, there is a 0.137 and 0.057 probability increase for the F/I injury in classes 1 and 4, respectively, while case 2 shows a −0.011 probability decrease for the F/I injury. For the bus-related case, heterogeneity results are also observed in the F/I injury, while the probability decrease for the N/P injury indicates an increase in the severe outcome in bus-related crashes. All these heterogeneities indicate a need to analyze the influences of these factors under the specific scenario, which again shows the superiority of using latent class random parameter models.

5.3.4. Locality and roadway characteristics

Compared to the rural area, crashes that occurred in the urban area also show heterogeneous results for the F/I injury. Similar conclusions on such heterogeneity were also made in Li and Fan (2019a). Debates on whether pedestrian-vehicle crashes happened in the urban area are safer than those in the rural area could be found in past studies. Some scholars concluded that rural is more dangerous because of the higher speed of vehicles and lack of medical resource (Sasidharan et al., 2015; Ulak et al., 2017), while others argued that urban has more complex traffic conditions with sufficiently high speed for fatality (Sun et al., 2019). These two explanations could well illustrate the heterogeneity in severe injuries that occurred in complex urban areas. In regard to land development, heterogeneous results could also be found in commercial land. Both commercial and institutional land show about a 0.04 probability increase for F/I injury in class 4, and class 4 denotes a latent class of urban local street without traffic control.

Roadway alignment, class level, and settlement are three major significant factors within the category of roadway features detected in this study. Compared to the straight-road, the curve-road shows a 0.046 and a 0.104 probability increase for the F/I injury in classes 3 and 4, respectively. Results on the grade-road show that the probability of pedestrians being F/I injured is increased by 0.023 and 0.076 compared to the level-road in classes 3 and 4, respectively. These locations are accident-prone areas as the driver has bad sight condition and the vehicle is difficult to control. Similar results could be referred to Sasidharan et al. (2015). Compared to the U.S. route, results on the interstate roads indicates a 0.158 and 0.046 probability increase for the F/I injury in classes 1 and 3. Also, the public vehicular area shows a −0.068 and −0.089 probability decrease for the F/I injury in classes 3 and 4. The reason for this might be that the public vehicular area (e.g., parking lot) has much lower traveling speeds than the U.S. route (Li & Fan, 2019b). Heterogeneities are also found in local streets and driveways. Classes 3 and 4 decrease the probability of pedestrians being F/I injured by −0.031 and −0.106, respectively, while class 2 shows a 0.038 increase in the F/I injury.

5.3.5. Environment and time characteristics

Compared to the daylight environment, both with/without lighting in the dark environment increase the probability of pedestrians being F/I injured in classes 1, 2, and 3 by 0.111, 0.197, and 0.061, respectively. The significant possibility decrease of the F/I injury requires a better lighting facility in these hotspots and this finding is in accordance with Yasmin et al. (2014). In comparison with the clear weather, the cloudy weather situation increases the probability of pedestrians being F/I injured by 0.073. A similar result could be referred to Aziz et al. (2013), and one possible reason for this is the decrease of sight in cloudy condition.

Though previous research has already pointed out the positive correlation between the vehicle-vehicle crash injury severity level with the peak hour (Mohamed et al., 2013), pedestrian-vehicle crash frequency in this study does not show a significant difference between the peak and non-peak hours as vehicle-vehicle crash does. Hence, this paper categorizes the crash time into six different periods mainly according to different features of the light condition, the frequency of the total crashes, and frequency of F/I injuries. Compared to the "morning" period (6:00–9:59), the "early morning" (0:00–5:59) shows a 0.053 probability increase for the F/I injury in class 4. Similar conclusions are showed in Haleem, Alluri, and Gan (2015). The "noon" (10:00–14:59), "afternoon" (15:00–17:59), and "early night" (21:00–23:59) in class 3 and class 4 all show a probability decrease of pedestrians being F/I injured. Furthermore, comparing the periods within class 3, results show that the afternoon hour has a higher probability of pedestrians being F/I injured than the noon hour. Result in class 4 indicates that pedestrian-vehicle crashes during the early night hour have a higher probability of being F/I injured for pedestrians compared to the afternoon hour. The increase of the F/I injury in the early morning might be caused by the combined impacts of dark environment, high speed, and fatigue of the driver in the early morning.

5.3.6. Traffic control characteristics

Compared to the situation of no traffic control, heterogeneity is found under the traffic sign control situation. Results on traffic sign control indicate a 0.016 probability decrease for the F/I injury in class 3, and a 0.028 probability increase for the F/I injury in class 4. The downward tendency of the probability of the NI injury in classes 3 and 4 is also detected. There are debates on the heterogeneous effects of traffic sign control on the safety of pedestrians. Kim et al. (2010) observed heterogeneity for traffic sign control in pedestrian fatalities, and the correlation between pedestrian age and traffic sign was detected. Possible explanations for such a difference in the effect of this factor could be concluded as follows: (a) the mitigatory outcome might result from the warning function of the traffic signs; and (b) the deteriorative result might be the consequence of the dangerous and complex environment where the traffic signs were installed.

6. Conclusions

This study explores factors of pedestrian-injury severity in pedestrian-vehicle crashes at hotspots with an upward trend considering the heterogeneity within and between the datasets. Twelve years of the police-reported pedestrian-vehicle crash data from 2007 to 2018 in North Carolina are used. Spatiotemporal trend analysis combined with the average nearest neighbor analysis and the spatial autocorrelation test are implemented to test the spatial clustering pattern and the temporal tendency of the crashes. The latent class clustering and four random parameter logit models are implemented to further investigate the heterogeneity within each class. Marginal effects are further calculated for better interpreting the impacts of categorical variables on the severity levels.

The random parameter variables detected across observations and the heterogeneous results between the subgroups indicate the superiority of combining the latent class clustering with random parameter logit models. Significant impacts of pedestrian behaviors, such as dash/dart-out and crossing or staying in the roadway, also require more attention to improve the transportation facilities to provide better protection for pedestrians. Meanwhile, there is a need to strengthen law enforcement and education to prohibit playing in roadways, crossing divided roadways without permission, and drunk walking in/across the roadways. Also, more appropriate traffic control management, such as adjusting the signal phase to decrease the behavior of crossing with the red light, is needed for both drivers and pedestrians. Besides, the zebra crossing sign could be equipped with flashing lights to alert the driver when the pedestrian is crossing since early night hour (0:00–5:59) is found to be the most dangerous period for pedestrians. Furthermore, a patrol route considering hotspots with an upward trend could help to reduce the response time to reach crash locations.

This paper provides a framework for researchers and engineers to identify crash hotspots considering spatiotemporal patterns and explore contribution factors to crashes considering unobserved heterogeneity. However, the temporal fluctuations may still exist in different time scales and may be caused by different factors such as the global recession (Behnood & Mannering, 2016). Further studies are still needed to investigate the heterogeneities within the time–space scale, spatial and temporal correlations of the factors, and the temporal fluctuation and instability of the crash data.

Acknowledgements

The authors want to express their deepest gratitude for the financial support from the United States Department of Transportation, University Transportation Center through the Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE) at The University of North Carolina at Charlotte (Grant Number: 69A3551747133).

References

Abay, K. A. (2013). Examining pedestrian-injury severity using alternative disaggregate models. *Research in Transportation Economics, 43*(1), 123–136. https://doi.org/10.1016/j.retrec.2012.12.002.

Anderson, T. K. (2009). Kernel density estimation and K-means clustering to profile road accident hotspots. *Accident Analysis & Prevention, 41*(3), 359–364. https://doi.org/10.1016/j.aap.2008.12.014.

Aziz, H. M. A., Ukkusuri, S. V., & Hasan, S. (2013). Exploring the determinants of pedestrian–vehicle crash severity in New York City. *Accident Analysis & Prevention, 50*, 1298–1309. https://doi.org/10.1016/j.aap.2012.09.034.

Behnood, A., & Mannering, F. L. (2016). An empirical assessment of the effects of economic recessions on pedestrian-injury crashes using mixed and latent-class models. *Analytic Methods in Accident Research, 12*, 1–17. https://doi.org/10.1016/j.amar.2016.07.002.

Blazquez, C. A., Picarte, B., Calderón, J. F., & Losada, F. (2018). Spatial autocorrelation analysis of cargo trucks on highway crashes in Chile. *Accident Analysis and Prevention, 120*(August), 195–210. https://doi.org/10.1016/j.aap.2018.08.022.

Chen, Z., & Fan, W. (. (2019). A multinomial logit model of pedestrian-vehicle crash severity in North Carolina. *International Journal of Transportation Science and Technology, 8*(1), 43–52. https://doi.org/10.1016/j.ijtst.2018.10.001.

Dai, D. (2012). Identifying clusters and risk factors of injuries in pedestrian–vehicle crashes in a GIS environment. *Journal of Transport Geography, 24*, 206–214. https://doi.org/10.1016/j.jtrangeo.2012.02.005.

Derr, B. (2013). Ordinal response modeling with the LOGISTIC procedure. *SAS Global Forum*, 1–20.

Ebdon, D. (1985). *Statistics in geography*. Blackwell.

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis, 24*(3), 189–206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x.

Gudes, O., Varhol, R., Sun, Q. (., & Meuleners, L. (2017). Investigating articulated heavy-vehicle crashes in western Australia using a spatial approach. *Accident Analysis & Prevention, 106*, 243–253. https://doi.org/10.1016/j.aap.2017.05.026.

Haleem, K., Alluri, P., & Gan, A. (2015). Analyzing pedestrian crash injury severity at signalized and non-signalized locations. *Accident Analysis & Prevention, 81*, 14–23. https://doi.org/10.1016/j.aap.2015.04.025.

Hu, L., Wu, X., Huang, J., Peng, Y., & Liu, W. (2020). Investigation of clusters and injuries in pedestrian crashes using GIS in Changsha, China. *Safety Science, 127*, 104710. https://doi.org/10.1016/j.ssci.2020.104710.

Iranitalab, A., & Khattak, A. (2017). Comparison of four statistical and machine learning methods for crash severity prediction. *Accident Analysis and Prevention, 108*(August), 27–36. https://doi.org/10.1016/j.aap.2017.08.008.

Kendall, M. G., Gibbons, J. D. (1990). Rank correlation methods (fifth ed.).

Kim, J.-K., Ulfarsson, G. F., Shankar, V. N., & Mannering, F. L. (2010). A note on modeling pedestrian-injury severity in motor-vehicle crashes with the mixed logit model. *Accident Analysis & Prevention, 42*(6), 1751–1758. https://doi.org/10.1016/j.aap.2010.04.016.

Lanza, S. T., Collins, L. M., Lemmon, D. R., & Schafer, J. L. (2007). PROC LCA: A SAS procedure for latent class analysis. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(4), 671–694. https://doi.org/10.1080/10705510701575602.

Li, D., Ranjitkar, P., Zhao, Y., Yi, H., & Rashidi, S. (2017). Analyzing pedestrian crash injury severity under different weather conditions. *Traffic Injury Prevention, 18*(4), 427–430. https://doi.org/10.1080/15389588.2016.1207762.

Li, Y., & Fan, W. (David) (2019a). Modelling severity of pedestrian-injury in pedestrian-vehicle crashes with latent class clustering and partial proportional odds model: A case study of North Carolina. *Accident Analysis and Prevention, 131*, 284–296. https://doi.org/10.1016/j.aap.2019.07.008.

Li, Y., & Fan, W. (David) (2019b). Pedestrian injury severities in pedestrian-vehicle crashes and the partial proportional odds logit model: Accounting for age difference. *Transportation Research Record, 2673*(5), 731–746. https://doi.org/10.1177/0361198119842828.

Liu, J., Hainen, A., Li, X., Nie, Q., & Nambisan, S. (2019). Pedestrian injury severity in motor vehicle crashes: An integrated spatio-temporal modeling approach. *Accident Analysis & Prevention, 132*, 105272. https://doi.org/10.1016/j.aap.2019.105272.

Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica, 13*(3), 245. https://doi.org/10.2307/1907187.

Mannering, F. (2018). Temporal instability and the analysis of highway accident data. *Analytic Methods in Accident Research, 17*, 1–13. https://doi.org/10.1016/j.amar.2017.10.002.

Mannering, F. L., & Bhat, C. R. (2014). Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research, 1*, 1–22. https://doi.org/10.1016/j.amar.2013.09.001.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics, 15*(5), 447–470. https://doi.org/10.1002/1099-1255(200009/10)15:5<447::AID-JAE570>3.0.CO;2-1.

McLachlan, G. J., & Peel, D. (2004). *Finite mixture models*. John Wiley & Sons. https://doi.org/10.1146/annurev-statistics-031017-100325.

Mohamed, M. G., Saunier, N., Miranda-Moreno, L. F., & Ukkusuri, S. V. (2013). A clustering regression approach: A comprehensive injury severity analysis of pedestrian–vehicle crashes in New York, US and Montreal, Canada. *Safety Science, 54*, 27–37. https://doi.org/10.1016/j.ssci.2012.11.001.

Mokhtarimousavi, S. (2019). A time of day analysis of pedestrian-involved crashes in California: Investigation of injury severity, a logistic regression and machine learning approach using HSIS data. *Institute of Transportation Engineers. ITE Journal, 89*(10), 25–33.

Mokhtarimousavi, S., Anderson, J. C., Azizinamini, A., & Hadi, M. (2020). Factors affecting injury severity in vehicle-pedestrian crashes: A day-of-week analysis using random parameter ordered response models and Artificial Neural Networks. *International Journal of Transportation Science and Technology*, 1–16. https://doi.org/10.1016/j.ijtst.2020.01.001.

Moran, P. A. P. (1948). The interpretation of statistical maps. In Journal of the Royal Statistical Society. Series B (Methodological) (Vol. 10, pp. 243–251). Wiley Royal Statistical Society. https://doi.org/10.2307/2983777.

Mujalli, R. O., Garach, L., López, G., & Al-Rousan, T. (2019). Evaluation of injury severity for pedestrian-vehicle crashes in Jordan using extracted rules. *Journal of Transportation Engineering Part A: Systems, 145*(7), 1–13. https://doi.org/10.1061/JTEPBS.0000244.

NHTSA. (2019). Traffic safety facts, 2017 data: Pedestrians. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812681.

Ouni, F., & Belloumi, M. (2018). Spatio-temporal pattern of vulnerable road user's collisions hot spots and related risk factors for injury severity in Tunisia. *Transportation Research Part F: Traffic Psychology and Behaviour, 56*, 477–495. https://doi.org/10.1016/j.trf.2018.05.003.

Plug, C., Xia, J., & Caulfield, C. (2011). Spatial and temporal visualisation techniques for crash analysis. *Accident Analysis & Prevention, 43*(6), 1937–1946. https://doi.org/10.1016/j.aap.2011.05.007.

Sasidharan, L., Wu, K. F., & Menendez, M. (2015). Exploring the application of latent class cluster analysis for investigating pedestrian crash injury severities in Switzerland. *Accident Analysis & Prevention, 85*, 219–228. https://doi.org/10.1016/j.aap.2015.09.020.

Song, L., & Fan, W. (2020). Combined latent class and partial proportional odds model approach to exploring the heterogeneities in truck-involved severities at cross and T-intersections. *Accident Analysis & Prevention, 144*, 105638. https://doi.org/10.1016/j.aap.2020.105638.

Songchitruksa, P., & Zeng, X. (2010). Getis-ord spatial statistics to identify hot spots by using incident management data. *Transportation Research Record, 2165*, 42–51. https://doi.org/10.3141/2165-05.

Sun, M., Sun, X., & Shan, D. (2019). Pedestrian crash analysis with latent class clustering method. *Accident Analysis & Prevention, 124*, 50–57. https://doi.org/10.1016/j.aap.2018.12.016.

Ulak, M. B., Kocatepe, A., Ozguven, E. E., Horner, M. W., & Spainhour, L. (2017). Geographic information system–based spatial and statistical analysis of severe crash hotspot accessibility to hospitals. *Transportation Research Record, 2635*(1), 90–97. https://doi.org/10.3141/2635-11.

Wang, C. W., & Chan, C. L. W. (2016). Estimated trends and patterns of road traffic fatalities in China, 2002–2012. *Traffic Injury Prevention, 17*(2), 164–169. https://doi.org/10.1080/15389588.2015.1045501.

Yalcin, G., & Sebnem Duzgun, H. S. (2015). Spatial analysis of two-wheeled vehicles traffic crashes: Osmaniye in Turkey. *KSCE Journal of Civil Engineering, 19*(7), 2225–2232. https://doi.org/10.1007/s12205-015-0661-0.

Yasmin, S., Eluru, N., & Ukkusuri, S. V. (2014). Alternative ordered response frameworks for examining pedestrian injury severity in New York City. *Journal of Transportation Safety & Security, 6*(4), 275–300. https://doi.org/10.1080/19439962.2013.839590.

Li Song received the M.S. degree in transportation engineering from the Harbin Institute of Technology. He is currently a Ph.D. Research Assistant of the USDOT Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE) at the Department of Civil and Environmental Engineering in University of North Carolina at Charlotte. His research interests include traffic safety and traffic control.

Wei (David) Fan received the Ph.D. degree in transportation engineering from the University of Texas at Austin. He is currently the Director of the USDOT Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE) and the Professor at the Department of Civil and Environmental Engineering in University of North Carolina at Charlotte. His research interests include big data analytics for transportation, connected and autonomous vehicles, multimodal transportation and shared mobility; traffic system operation and control, and transportation system analysis and network modeling.

Yang Li received the M.S. degree in transportation engineering from the Clemson University. He is currently a Ph.D. Research Assistant in the USDOT Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE) at the Department of Civil and Environmental Engineering in University of North Carolina at Charlotte.His research interests include traffic safety and data analysis.

Peijie Wu received the B.S. degree in transportation engineering from Chongqing Jiaotong University. She is currently a Ph.D. student at the School of Transportation Science and Engineering in Harbin Institute of Technology. Her research interests include traffic safety and data analysis.

Appendix A

Table A1
Statistics of explanatory variables for pedestrian-vehicle crashes at hotspots with an upward trend.

Variable	Description		Total	F/I[a]	NI[b]	N/P[c]
	Number of observations		13303	1415(10.64%)	5046(37.93%)	6842(51.43%)
Pedestrian Characteristics						
PedAge	**PedAge ≤ 24**	1	4305	395(9.18%)	1788(41.53%)	2122(49.29%)
	24 < PedAge ≤ 54	2	6459	680(10.53%)	2330(36.07%)	3449(53.4%)
	55 < PedAge ≤ 64	3	1299	164(12.63%)	449(34.57%)	686(52.81%)
	PedAge ≥ 65	4	1240	176(14.19%)	479(38.63%)	585(47.18%)
PedSex	**Male**	1	7593	972(12.8%)	2978(39.22%)	3643(47.98%)
	Female	2	5710	443(7.76%)	2068(36.22%)	3199(56.02%)
PedAlcFlag	**PedAlcFlag = 'No'**	1	11867	1026(8.65%)	4416(37.21%)	6425(54.14%)
	PedAlcFlag = 'Yes'	2	1436	389(27.09%)	630(43.87%)	417(29.04%)
Vehicle Type						
DrvrVehTyp	**Small**	1	7998	754(9.43%)	3045(38.07%)	4199(52.5%)
	Middle	2	4940	593(12%)	1865(37.75%)	2482(50.24%)
	Heavy	3	365	68(18.63%)	136(37.26%)	161(44.11%)
Crash Characteristics						
AmbulanceR	**Ambulance Rescue**	1	9880	1287(13.03%)	4225(42.76%)	4368(44.21%)
	No Ambulance Rescue	2	3423	128(3.74%)	821(23.98%)	2474(72.28%)
HitRun	**No Hit and Run**	1	11653	1279(10.98%)	4494(38.57%)	5880(50.46%)
	Hit and Run	2	1650	136(8.24%)	552(33.45%)	962(58.3%)
CrashGrp	**Walking Along Roadway**	1	845	117(13.85%)	346(40.95%)	382(45.21%)
	Crossing Roadway with Vehicle Not Turning	2	2692	502(18.65%)	1079(40.08%)	1111(41.27%)
	Crossing Roadway with Vehicle Turning	3	2092	60(2.87%)	718(34.32%)	1314(62.81%)
	Off Roadway	4	1632	59(3.62%)	486(29.78%)	1087(66.61%)
	Pedestrian in Roadway	5	696	135(19.4%)	256(36.78%)	305(43.82%)
	Dash/Dart-Out	6	1182	175(14.81%)	632(53.47%)	375(31.73%)
	Backing Vehicle	7	1325	54(4.08%)	342(25.81%)	929(70.11%)
	Multiple Threat/Trapped	8	214	15(7.01%)	113(52.8%)	86(40.19%)
	Bus related Vehicle	9	134	12(8.96%)	64(47.76%)	58(43.28%)
	Other/Unusual Circumstances	10	2491	286(11.48%)	1010(40.55%)	1195(47.97%)
Locality and Roadway Characteristics						
Locality	**Rural**	1	1282	276(21.53%)	485(37.83%)	521(40.64%)
	Urban	2	12021	1139(9.48%)	4561(37.94%)	6321(52.58%)
Development	**Residential**	1	4584	484(10.56%)	1897(41.38%)	2203(48.06%)
	Commercial	2	7705	768(9.97%)	2769(35.94%)	4168(54.09%)
	Industrial	3	76	3(3.95%)	37(48.68%)	36(47.37%)
	Institutional	4	485	27(5.57%)	172(35.46%)	286(58.97%)
	Farms, Woods, Pastures	5	453	133(29.36%)	171(37.75%)	149(32.89%)
RdCurve	**Straight**	1	12770	1306(10.23%)	4846(37.95%)	6618(51.82%)
	Curve	2	533	109(20.45%)	200(37.52%)	224(42.03%)
RdGrad	Level	1	10996	1055(9.59%)	4116(37.43%)	5825(52.97%)
	Grade	2	1718	277(16.12%)	685(39.87%)	756(44%)
	Hillcrest	3	502	66(13.15%)	203(40.44%)	233(46.41%)
	Bottom	4	87	17(19.54%)	42(48.28%)	28(32.18%)
RdClass	**US Route**	1	415	128(30.84%)	170(40.96%)	117(28.19%)
	Interstate	2	241	103(42.74%)	72(29.88%)	66(27.39%)
	State Route	3	329	78(23.71%)	138(41.95%)	113(34.35%)
	State Secondary Route	4	465	100(21.51%)	196(42.15%)	169(36.34%)
	Local Street, Driveway	5	8479	882(10.4%)	3521(41.53%)	4076(48.07%)
	Public Vehicular Area	6	3374	124(3.68%)	949(28.13%)	2301(68.2%)

(continued on next page)

Appendix A (*continued*)

Variable	Description		Total	F/I[a]	NI[b]	N/P[c]
RdConfig	**One-Way, Not Divided**	1	1275	69(5.41%)	421(33.02%)	785(61.57%)
	Two-Way, Not Divided	2	9119	820(8.99%)	3395(37.23%)	4904(53.78%)
	Two-Way, Divided	3	2909	526(18.08%)	1230(42.28%)	1153(39.64%)
Environment and Temporal Characteristics						
LightCond	**Daylight**	1	7977	519(6.51%)	2923(36.64%)	4535(56.85%)
	Dawn/Dusk Light	2	604	62(10.26%)	207(34.27%)	335(55.46%)
	Dark – Lighted Roadway	3	3336	456(13.67%)	1390(41.67%)	1490(44.66%)
	Dark – Roadway Not Lighted	4	1386	378(27.27%)	526(37.95%)	482(34.78%)
Weather	**Clear**	1	10242	1075(10.5%)	3909(38.17%)	5258(51.34%)
	Cloudy	2	1800	204(11.33%)	643(35.72%)	953(52.94%)
	Rain	3	1141	125(10.96%)	440(38.56%)	576(50.48%)
	Snow, Sleet, Hail, Freezing Rain/Drizzle	4	80	4(5%)	38(47.5%)	38(47.5%)
	Fog, Smog, Smoke	5	40	7(17.5%)	16(40%)	17(42.5%)
Hour	**6:00–9:59**	1	1889	174(9.21%)	709(37.53%)	1006(53.26%)
	10:00–14:59	2	3252	199(6.12%)	1107(34.04%)	1946(59.84%)
	15:00–17:59	3	2805	200(7.13%)	1064(37.93%)	1541(54.94%)
	18:00–20:59	4	2668	330(12.37%)	1036(38.83%)	1302(48.8%)
	21:00–23:59	5	1598	269(16.83%)	671(41.99%)	658(41.18%)
	0:00–5:59	6	1091	243(22.27%)	459(42.07%)	389(35.66%)
Traffic Control Type						
TraffCntrl	**No Control Present**	1	8876	996(11.22%)	3351(37.75%)	4529(51.03%)
	Signs	2	1126	73(6.48%)	367(32.59%)	686(60.92%)
	Signal	3	2613	220(8.42%)	1056(40.41%)	1337(51.17%)
	Double Yellow Line, No Passing Zone	4	548	119(21.72%)	230(41.97%)	199(36.31%)
	Human Control	5	140	7(5%)	42(30%)	91(65%)

Note: variables in bold and numbered with 1 are set as the base for the explanatory variables.
[a] F/I – Fatal/Incapacitating injury. [b]NI – Non-incapacitating injury. [c] N/P – No/Possible injury.

Table A2
Random parameter logit model's significant variable coefficients for class 1.

Variable	Description	F/I[a]		NI[b]	
		Coef.	t value	Coef.	t value
Intercept		−1.0814	−5.05	0.3277	3.11
PedAge4	PedAge ≥ 65	1.2747	3.68		
PedAlcFlag2	PedAlcFlag = 'Yes'	0.4824	1.92		
DrvrVehTyp3	Heavy	1.2946	2.97		
HitRun2	Hit and Run	−0.8608	−1.89		
AmbulanceR2	No ambulance rescue	−2.2402	−5.24	−1.5076	−6.97
CrashGrp2	Crossing roadway with vehicle not turning	0.9042	3.32		
CrashGrp6	Dash/Dart-Out	1.2366	3.46		
CrashGrp7	Backing vehicle			−2.2548	−2
Locality2	Urban	−0.6561	−1.8		
RdGrad2	Grade			0.4057	2.11
RdClass2	Interstate	0.8898	2.41		
RdConfig3	Two-Way, Divided	1.1638	3.57	0.5615	2.52
LightCond4	Dark – Roadway Not Lighted	0.7111	2.85		
Std. dev.		1.2284	1.81		

Note: Number of observations: 901. Log-likelihood at convergence: −862.41. Log-likelihood (constant only): −989.85.
[a] F/I – Fatal/Incapacitating injury. [b] NI – Non-incapacitating injury.

Table A3
Random parameter logit model's significant variable coefficients for class 2.

Variable	Description	F/I[a]		NI[b]	
		Coef.	t value	Coef.	t value
Intercept		−1.9696	−5.98	−1.395	−4.94
PedAge2	24 < PedAge ≤ 54	−0.6698	−3.34	−0.309	−3.09
PedAge4	PedAge ≥ 65			0.4966	3.84
DrvrVehTyp2	Middle	0.681	3.41	0.2114	2.34
DrvrVehTyp3	Heavy	1.9061	4.97	0.6345	2.61
AmbulanceR2	No Ambulance Rescue	−1.9317	−6.46	−0.9216	−8.56
CrashGrp4	Off Roadway			0.7939	2.35
Std. dev.				1.3123	2.3
CrashGrp7	Backing Vehicle			0.8114	3
CrashGrp10	Other/Unusual Circumstances	0.9622	4.6	1.3561	4.97
Locality2	Urban	−0.5765	−2.05		
Development2	Commercial	−0.5448	−2.63	−0.4106	−4.09
RdClass5	Local Street, Driveway	0.9522	3.09	0.3767	2.32
Weather2	Cloudy	−0.5732	−1.79		
Hour6	0:00–5:59	1.0055	2.92	0.7169	3.92
TraffCntrl4	Double Yellow Line, No Passing Zone	2.463	1.95		

Note: Number of observations: 3517. Log-likelihood at convergence: −2432. Log-likelihood (constant only): −3864.

Table A4
Random parameter logit model's significant variable coefficients for class 3.

Variable	Description	F/I[a]		NI[b]	
		Coef.	t value	Coef.	t value
Intercept		−0.7485	−3.1	1.4015	4.23
PedAge3	55 < PedAge ≤ 64	0.4373	2.48		
PedAge4	PedAge ≥ 65	0.9448	5.5		
PedAlcFlag2	PedAlcFlag = 'Yes'	1.2401	5.04	0.6698	2.3
DrvrVehTyp2	Middle	0.4965	4.23		
HitRun2	Hit and Run	−0.6081	−2.67	−0.4504	−2.4
AmbulanceR2	No Ambulance Rescue	−1.8857	−8.92	−1.5342	−5
CrashGrp3	Crossing Roadway with Vehicle Turning	−1.7686	−9.54	−0.5091	−3.74
CrashGrp6	Dash/Dart-Out			0.9431	3.69
CrashGrp8	Multiple Threat/Trapped			0.7118	2.47
CrashGrp9	Bus related Vehicle			0.973	2.3
Locality2	Urban			−0.7269	−2.86
Std. dev.				2.1824	2.86
Development5	Farms, Woods, Pastures	3.7492	3.47	2.2967	1.71
RdCurve2	Curve	0.5734	2.5		
RdGrad2	Grade	0.3173	2.17		
RdClass2	Interstate			−1.9342	−2.84
RdClass5	Local Street, Driveway	−0.8247	−3.69	−0.7765	−2.96
RdClass6	Public Vehicular Area	−2.7034	−2.64	−1.875	−2.72
RdConfig3	Two-Way, Divided	0.2928	2.29		
LightCond2	Dawn/Dusk Light			−0.3717	−1.66
LightCond3	Dark – Lighted Roadway	1.728	4.33		
Weather4	Snow, Sleet, Hail, Freezing Rain/Drizzle			1.6402	1.92
Hour2	10:00–14:59	−0.4597	−3.21		
Hour3	15:00–17:59	−0.2804	−1.98		
TraffCntrl2	Signs	−0.4593	−2.24	−0.4451	−2.57
TraffCntrl4	Double Yellow Line, No Passing Zone	0.9243	2.56		

Note: Number of observations: 5255. Log-likelihood at convergence: −4358. Log-likelihood (constant only): −5773.

Table A5
Random parameter logit model's significant variable coefficients for class 4.

Variable	Description	F/I[a]		NI[b]	
		Coef.	t value	Coef.	t value
Intercept		−1.7074	−8.05	0.3974	3.75
PedAge2	24<PedAge ≤ 54	0.28	2.27		
PedAge3	55<PedAge ≤ 64	0.6125	3.44		
PedAge4	PedAge ≥ 65	1.1367	5.39		
PedAlcFlag2	PedAlcFlag = 'Yes'	0.7554	6.28	0.365	3.1
Std. dev.				1.3175	1.93
DrvrVehTyp3	Heavy	1.0348	2.8		
HitRun2	Hit and Run			−0.4551	−2.08
Std. dev.				1.5665	1.88
AmbulanceR2	No Ambulance Rescue	−1.3145	−8.11	−1.1283	−9.9
CrashGrp2	Crossing Roadway with Vehicle Not Turning	0.8225	5.94	0.2264	2.21
CrashGrp3	Crossing Roadway with Vehicle Turning	−1.4071	−5.24	−0.6261	−4.98
CrashGrp5	Pedestrian in Roadway	0.7605	4.16		
CrashGrp6	Dash/Dart-Out	0.9161	4.9	0.4135	2.88
CrashGrp7	Backing Vehicle			−0.843	−2.81
CrashGrp9	Bus related Vehicle	1.317	2.17		
Development2	Commercial	0.3504	3.08		
Development4	Institutional			−0.6541	−2.42
RdCurve2	Curve	0.717	3.13		
RdGrad2	Grade	0.5511	4.07		
RdClass5	Local Street, Driveway	−0.7426	−5.06		
RdClass6	Public Vehicular Area	−1.2944	−3.36	−0.7613	−2.59
RdConfig2	Two-Way, Not Divided			−0.3195	−3.71
RdConfig3	Two-Way, Divided	0.7398	6.45		
LightCond4	Dark – Roadway Not Lighted	0.4588	3.98		
Weather2	Cloudy	0.534	3.74		
Hour3	15:00–17:59	−0.9898	−3.06		
Hour5	21:00–23:59			0.2262	2.56
Hour6	0:00–5:59	0.6473	5.02	0.3936	3.34
TraffCntrl2	Signs			−0.3854	−2.59
TraffCntrl4	Double Yellow Line, No Passing Zone	0.6512	1.99		

Note: Number of observations: 3630. Log-likelihood at convergence: −3361. Log-likelihood (constant only): −3988.

Journal of Safety Research 76 (2021) 269–275

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Factors affecting pedestrian behaviors at signalized crosswalks: An empirical study

Salaheddine Bendak [a,*], Asayel M. Alnaqbi [b], Muna Y. Alzarooni [b], Sara M. Aljanaahi [b], Shaikha J. Alsuwaidi [b]

[a] *Department of Industrial Engineering, Halic University, Istanbul, Turkey*
[b] *Department of Industrial Engineering and Engineering Management, University of Sharjah, Sharjah, United Arab Emirates*

ARTICLE INFO

Article history:
Received 1 April 2020
Received in revised form 24 September 2020
Accepted 21 December 2020
Available online 7 January 2021

Keywords:
Pedestrian behavior
Road safety
Crosswalks
Pedestrian crashes
Health Belief Model

ABSTRACT

Introduction: Safety of pedestrians depends, among other factors, on their behavior while crossing the road. This study aims to assess behaviors of pedestrians at signalized crosswalks. *Method:* Following a literature review and a pilot study, 25 vital pedestrian crossing factors and behaviors were determined. Then data was randomly collected for 708 pedestrians at 10 lighted crossings in Sharjah (UAE), five at road intersections and five mid-block crossings. *Results:* Results indicated that 17.4% of pedestrians observed crossed partly or fully on red and that crossing speed was 1.22 m/s, on the average, which is slightly faster than most speeds recorded in the literature. Moreover, female pedestrians were more likely to cross while chatting with others, less likely to cross on red, and more likely to walk slower than male pedestrians. Results also showed that pedestrians who crossed at road intersections walked slower than those who crossed at mid-block crossings. It was also found that longer red pedestrian times and narrower roads tended to encourage pedestrians to cross on red and that the majority of pedestrians did not look around before crossing. *Practical implications*: Use of the Health Belief Model for pedestrian safety are discussed.

1. Introduction

Of the 1.35 million deaths due to road traffic collisions annually, nearly one-quarter (23%) are pedestrians. Traffic crashes involving pedestrians constitute a major road safety problem in the entire world, especially in developing countries. Pedestrians are among the most often neglected groups in many countries when designing road traffic systems. This can mainly be attributed to the rapid increase in population, increase in the number of motor vehicles in urban centers, and to disobedience of traffic regulations by pedestrians, drivers and other road users. Therefore, there is a dire need to understand risk factors that affect the safety of pedestrians on the roads (Alhajyaseen & Iryo-Asano, 2017; Choi et al., 2019; Jain et al., 2014; Tezcan et al., 2019; WHO, 2018; Zegeer & Bushell, 2012).

Signalized crosswalks are designated walking paths for pedestrians with red-green signal lights. They are widely used on roads where traffic volume is high and involve stopping motor vehicles to allow pedestrians to cross safely from one side of the road to the other side. These signalized crossings usually are located at road junctions or at mid-blocks (i.e., at middle section of the road). Although these crossings are meant to serve as a measure allowing pedestrians and drivers safe passage, disobeying them can lead to the opposite. In other words, if pedestrians or drivers opt to violate pedestrian crossing signals for any reason, these crossings can lead to crashes.

A glance at the published scientific literature indicates that many studies stressed on the importance of microanalysis of pedestrian behaviors on crosswalks as a way to improve traffic safety (Alhajyaseen & Iryo-Asano, 2017; Koh et al., 2014; Papadimitriou et al., 2013; Tezcan et al., 2019). For example, Papadimitriou et al. (2013) reported that when a pedestrian is in a hurry and/or it is observed as harmless to unlawfully cross a road (i.e., there is slight or no contradictory traffic), he or she feels that they have an upper hand affinity to cross. In other instances, it might be as a result of following the crowd where the pedestrian just follows the crowd who is crossing unlawfully. The authors also reported that male pedestrians are further likely to begin unlawfully as compared to female pedestrians while mature pedestrians have the uppermost proportion of unlawful crossers when compared to young and elderly individuals. Overall, the authors found

* Corresponding author.
E-mail address: sbendak@halic.edu.tr (S. Bendak).

https://doi.org/10.1016/j.jsr.2020.12.019

that female pedestrians were more compliant with traffic regulations and that they exhibited more positive attitude and behaviors in relation to pedestrian traffic regulations than male pedestrians. The authors concluded that the road safety of pedestrians depends on their behavior, which is based on their beliefs, culture, perceptions and attitude.

In line with this last conclusion by Papadimitriou et al. (2013); Nordfjærn and Şimşekoğlu (2013) analyzed cultural and social-cognitive factors of risky pedestrian behaviors among pedestrians in Turkey using the Hofstede theory. The authors stated that this theory defines culture through its consequences for attitude, values and beliefs. They used uncertainty avoidance in Hofstede theory, defined as the level at which a person feels uncomfortable with ambiguity, to explain pedestrian attitude and behaviors when crossing roads. Moreover, they found that collectivism (which emphasizes norms) and individualism (which emphasizes attitude) are relevant in explaining pedestrian behaviors when crossing roads because these two cultural dimensions influence individual's conformity with regulations set by authorities. Uncertainty avoidance was also found to be relevant in evaluating pedestrian road crossing behaviors because people who score high in the cultural dimension are more risk-aversive, hence less likely to display risky road crossing behaviors. Their results showed also that vertical collectivism is associated with a low level of risk when crossing roads, while horizontal collectivism is related to greater risks when crossing roads.

Mwakalonge et al. (2015) investigated the effect of distractions on road crossing behaviors of pedestrians. The authors postulated that distractions affecting pedestrians at uncontrolled crossings in the United States root from their engagements in multitasking activities such as snacking, using hand-held devices, reading and listening to music. The authors found that out of 1,102 pedestrians, 29.8% showed distractive activities when crossing roads and concluded that distractive activities are related to pedestrian behaviors. Jain et al. (2014) reported similar conclusions in India and added that uncontrolled pedestrian crossings require high attention because pedestrians need to cross roads at a speed faster than vehicular movements.

The current study is conducted in the city of Sharjah, United Arab Emirates (UAE). Similar to the rest of the world, crashes involving pedestrians form a major road safety problem in UAE where the population exceeds nine million and the number of motor vehicles is around three million (FCSA-UAE, 2017). Number of pedestrians that died or got injured in UAE due to traffic crashes over three years is given in Table 1 (FCSA-UAE, 2017). As can be seen in that table, the average percentage of pedestrians killed among all traffic related fatalities is 24.2%, which is slightly higher than the world average of 23% (WHO, 2018).

Although the scientific literature included studies that aimed to capture pedestrian behaviors while crossing roads, most of those studies focused on specific behaviors. For Example, Koh et al. (2014) focused on analyzing violations at signalized crosswalks while Thompson et al. (2013) focused on capturing the effects of social and technological distractions on road crossing behaviors among pedestrians. Accordingly and as suggested by Mwakalonge et al. (2015) and Thompson et al. (2013), this study aims to capture and analyze a multitude of pedestrian behaviors at signalized crosswalks to provide a comprehensive and empirical view of the problem. It is anticipated that results of the current study would help in determining factors that shape the behavior of pedestrians at signalized crosswalks and be useful in managing and designing signalized crosswalks so that the safety of pedestrians would be enhanced.

2. Methodology

Recorded methods of pedestrian behavior on the road in the published scientific literature include direct roadside observations, questionnaire surveys, pedestrian crash analysis, or a mixture of them (Choi et al., 2019; Jain et al., 2014; Mwakalonge et al., 2015; Papadimitriou et al., 2013; Tezcan et al., 2019). It was thought in the current study that direct unobstructive roadside observations would be a reliable and objective way for data collection knowing that detailed data on pedestrian crashes are not available in UAE.

2.1. Data collection

There is a wide range of behaviors of pedestrians that can be observed at crosswalks. As discussed earlier and based on previous studies (specifically Jain et al., 2014; Koh et al., 2014; Mwakalonge et al., 2015; Thompson et al., 2018) a wide and comprehensive array of pedestrian behaviors, socio-demographic factors, distractive factors and other relevant road details were compiled and captured in the current study as follows:

1. Pedestrian gender
2. Day of week (weekday; weekend) (weekends consist of Thursday evenings, Fridays and Saturdays)
3. Number of pedestrians waiting for the green light at the curb (5 or less; 6 or more)
4. Ambient temperature (°C)
5. Type of crosswalk (road intersection; in-block)
6. Green time (sec)
7. Red time (sec)
8. Cycle time (sec)
9. Number of lanes each way
10. Age group (years) (less than 15; 16–39; 40 or more)
11. Carrying a load (yes/no)
12. Child walking with adults (yes/no)
13. Walking with a pedestrian child (yes/no)
14. Carrying a child (yes/no)
15. Pedestrian pushing pram or wheelchair (yes/no)
16. Riding a bicycle (yes/no)
17. Talking over mobile phone (handheld) (yes/no)
18. Talking over mobile phone (earphones) (yes/no)
19. Text messaging (yes/no)
20. Walking inside pedestrian designated area (yes/no)
21. Chatting with other pedestrians while walking (yes/no)
22. Crossing duration (s)
23. Crossing speed (m/s)
24. Crossing on red (yes/no)
25. Looking left and right prior to crossing (yes/no)

Table 1
Number of injuries and deaths due to road traffic crashes in UAE (FCSA-UAE, 2017).

Year	Total no. of injuries	No. of pedestrian injuries	Total no. of deaths	No. of pedestrian deaths
2014	7108	1084	712	174
2015	6865	995	675	157
2016	6681	1061	706	176

The first 16 factors were considered as input variables in the statistical analysis while the remaining nine factors reflected pedestrian behaviors while crossing.

Based on information solicited from Sharjah Road Traffic Authority, there are 38 signalized junctions with (red-green) signalized crosswalks and 46 in-block signalized crosswalks in Sharjah. In this study, it was decided to collect data at five signalized crosswalks at road intersections and at five in-block signalized crosswalks. Some of these 84 crossings were barely used based on prior knowledge and initial observations of the researchers. Accordingly, the 10 crosswalks chosen for this study were selected semi-randomly while making sure that the flow volume of pedestrians crossing the road is relatively high. Also, the timing of observations was decided to coincide with high pedestrian flow rate (usually 5–8 p.m. in Sharjah). It should be noted also that all of the 10 chosen crosswalks were on two-way, busy and, relatively speaking, wide streets with medians where pedestrian curbs were at least three meters wide on both sides.

Similar to some other studies like that of Choi et al. (2019), this study utilized unobtrusive video recording of pedestrians to analyze their crossing behaviors. Crossing pedestrians were video recorded in the following road intersections:

1. Institute of Banking junction (King Abdul Aziz St)
2. Gold Souq junction (University City Rd)
3. End of Esteqlal Rd (University Bookshop)
4. Dubai Islamic Bank Junction (King Abdul Aziz St)
5. Al-Intifadah Road and Al-Khan St junction

They were also recorded at the following in-block signalized crosswalks:

1. Al-Hilal Bank (after Institute of Banking) (King Abdul Aziz St)
2. Al-Zahra Hospital (Al-Zahra Rd roundabout) (Al-Zahra St)
3. Al-Majaz Theatre (Al-Buhaira) (Corniche St)
4. Pizza Hut Al-Buhaira (Corniche St)
5. Al-Intifadah Road (Near Al-Buhaira police station)

The 10 crosswalks are pinpointed on the Google Earth map given in Fig. 1. It was decided to collect data on pedestrians crossing over two to three traffic signal cycles at peak times on three different days at each of the above listed crosswalks. Two traffic signal cycles were recorded if the cycle duration was more than 200 s and three cycles were recorded if the duration was less than 200 s. Such a data volume (10 crosswalks X 2 (or 3) cycles X 3 repetitions) was anticipated to generate several hundred data points (based on previous studies like Jain et al. (2014) and Papadimitriou et al. (2013) and on initial observations by the researchers), which was thought to be a reasonable sample size. It should be noted that pedestrian crossing time countdown displays are not used in Sharjah.

2.2. Data analysis

The null hypothesis to be tested for all independent variables was that the medians of treatments of a given output variable across all levels of a given input variable are the same. Due to categorical nature of some of the variables (namely using mobile phone while walking, walking inside designated area, chatting with other pedestrians while walking, crossing on red, looking right and left prior to crossing and riding a bicycle), it was decided to use chi-square test for these variables. Moreover, normality test was done on crossing (walking) speed (which was the only continuous output variable) to determine the most suitable statistical analysis option. All statistical analyses were done using SPSS version 24.

3. Results and discussion

A pilot study was done at the beginning to monitor pedestrian behaviors at three pedestrian lights in Sharjah. The aim of this pilot study was to make sure that data on all important pedestrian behaviors would be recorded. Pilot study results indicated that many cyclists cross the street with pedestrians and that the percentage of people talking over the mobile phone (with either mode) or texting while walking was small. It was decided, therefore, to merge all modes of mobile phone usage into one called "crossing while using mobile phone" and to add a behavior named "crossing on a bicycle." Then data for the main study were collected.

In total, behaviors of 708 pedestrians from 30 recordings (three at each location) were recorded. Statistics on recorded temperature, crossing duration, green light time, red light time, and cycle time are given in Table 2. As can be seen in Table 2, the average walking speed at lighted intersections (calculated by dividing distance by duration) was found to be 1.22 m/s. Such a speed is slightly considered faster than most speeds reported in the literature (Goh et al., 2012). This might be due to the fact that close to 85% of residents in UAE are expatriates who reside in the country for work and, thus, tend to be younger than most other communities (FCSA, 2017).

In Table 3, more descriptive statistics of the data recorded are given. Following Table 3, statistical analysis results and discussion are given. It should be noted that no motor vehicle was recorded crossing when pedestrian lights were green throughout the whole study.

3.1. Gender effects

Male and female pedestrians were found to be different in terms of chatting while crossing the road ($p = 0.01$). It was recorded that 20.4% of observed male pedestrians crossed the road while chatting, while 30% of females were recorded chatting while crossing. They were also found different in crossing on red pedestrian lights ($p = 0.04$) where males (19.2%) tended to cross partly or fully on red pedestrian lights more often than females (13%). Possible reasons for this difference might include male pedestrians' greater willingness to take risk in crossing the road on red pedestrian lights and their greater ability, on the average, to walk faster than female pedestrians.

Crossing while riding a bicycle was also significantly different between the two genders ($p = 0.00$): 13.4% of males crossing were riding a bicycle, while only 1.9% of females were riding a bicycle. Females were also found to cross the road slower than males ($p = 0.00$), where the walking speed of female pedestrians was found to be 1.1 m/s on the average while that of male pedestrians was 1.27 m/s.

3.2. Age effects

Pedestrians between 16 and 39 years of age were found to significantly use their mobile phones (25.7%) while crossing ($p = 0.03$) compared to other age groups (9.8% for young pedestrians and 14.6% for older ones). Moreover, pedestrians between 16 and 39 years of age were found to significantly cross the road (partly or fully) on red more often than other age groups ($p = 0.00$), as the vast majority (94.3%) of the 123 pedestrians who crossed on red were in that age group.

3.3. Day of week effects

Looking right and left before crossing was found to be affected by day of week ($p = 0.04$), where 31.1% of pedestrians were

Fig. 1. Location of pedestrian crosswalks on Google Earth map.

Table 2
Descriptive statistics of temperature and crossing and light durations.

Statistic	Temp (°C)	Crossing duration (s)	Crossing speed (m/s)	Green time (s)	Red time (s)	Cycle time (s)
Mean	33.30	8.67	1.22	77.89	87.85	165.61
Std. Dev.	3.35	2.86	0.45	67.48	47.82	92.74
Minimum	27	2.19	0.64	15	27	45
Maximum	38	21.20	4.38	182	233	373

observed looking around before crossing on weekdays while 21.1% looked around over weekends. One possible reason for this difference is that pedestrians feel that looking around before crossing the road is less necessary on weekends when more pedestrians cross the road than on weekdays.

Walking speed was also found to be significantly different between weekdays and weekends ($p = 0.00$) where the average walking speed on weekdays was 1.26 m/s while it was 1.12 m/s on weekends. A likely reason for this difference is that on weekends, more families with children were observed crossing the roads than on weekdays.

3.4. Number of people waiting effects

Walking speed was also found to be significantly affected by the number of people crossing ($p = 0.03$). Pedestrians crossing with six or more other individuals walked at a speed of 1.15 m/s, while those with five or less pedestrians walked with an average speed

Table 3
Descriptive statistics of data recorded.

Factor	Level 1	Level 2	Level 3
Gender	male 501 (70.8%)	female 207 (29.2%)	
Age	≤15 61 (8.6%)	16 to 39 544 (76.8%)	≥40 103 (14.5%)
Day of week	weekend 123 (17.4%)	weekday 585 (82.6%)	
Number of lanes each way	2 lanes 72 (10.2%)	3 lanes 592 (83.6%)	4 lanes 44 (6.2%)
Number of pedestrians waiting	≤5 367 (51.8%)	≥6 341 (48.2%)	
Using mobile phone	talking (handheld) 24 (3.4%)	talking (earphones) 5 (0.1%)	texting 27 (3.8%)
Place of crossing	inside designated area 569 (80.4%)	(partly or fully) outside designated area 139 (19.6%)	
Crossing on red	Crossed on green 585 (82.6%)	crossed (partly or fully) on red 123 (17.4%)	
Crossing while chatting with others	not chatting 544 (76.8%)	chatting 164 (23.2%)	
Pedestrians walking with children/adults	adult walking with children 29 (4.1%)	adult carrying a child 4 (0.1%)	child walking with adults 31 (4.4%)
Pedestrian carrying a load, pushing others or riding a bicycle	carrying load 63 (8.9%)	riding bicycle 71 (10%)	pushing a pram or a wheelchair 13 (1.8%)

of 1.27 m/s. In other words, it was observed that pedestrians crossing with bigger crowds tend to walk slower than those crossing with smaller groups of pedestrians or alone.

3.5. Effects of type of pedestrian lights

Looking right and left before crossing was found to be significantly higher at mid-block pedestrian crosswalks (34.3%) than at road intersection pedestrian crosswalks (20.2%) (p = 0.00). Another difference was recorded in walking speed. Pedestrians who cross the street at road intersections walked slower (1.14 m/s) than those who crossed at mid-block signalized pedestrian lights (1.3 m/s). One likely reason for both results is that the green time at road junctions was longer, on the average, than that at mid-block signalized pedestrian lights, which might be giving more sense of security and leading to slower walking pace.

Using mobile phones while walking from one side of the road to the other was significantly less at junction pedestrian crosswalks (2.4%) than at mid-block crosswalks (10.6%) (p = 0.00). No meaningful reason for this phenomenon could be postulated.

3.6. Effects of walking with children

As can be anticipated, no pedestrian walking with a child (out of 29) was observed using a mobile phone while crossing the road while 8.1% of others (i.e., pedestrians not walking with children) were using their mobile phone while crossing. This difference was statistically significant (p = 0.04) and is believed to be the result of usually using a hand for holding a child's hand while walking, which makes it less likely for people walking with children to use their mobile phones, especially while crossing the road.

Pedestrians walking with child(ren) crossed the road, as expected, at slower speeds than others walking without children (p = 0.00). The average walking speed was 1.27 m/s for pedestrians crossing without children and 0.83 m/s for those accompanying children.

3.7. Effects of temperature

Pedestrians crossing the road were significantly less likely to use a mobile phone (p = 0.01), more likely to cross on red (p = 0.00) and more likely to walk faster (p = 0.00) at higher temperatures than at lower temperatures. Such an outcome is expected as people, in general, tend to do anything that would shorten their exposure to hot weather conditions.

3.8. Effects of green light duration

Pedestrians tend to walk significantly more often outside designated areas at crosswalks with shorter green times (p = 0.00) and longer red times (p = 0.00) than at other crosswalks. They also tend to cross on red significantly more often at crosswalks with shorter green times (p = 0.00) and longer red times (p = 0.00) than at other crosswalks. Moreover, pedestrians tend to walk significantly faster at crosswalks with shorter green times than at those with longer times (p = 0.00). Such behaviors may be attributed to pedestrians' less willingness to wait longer times at crosswalks.

3.9. Effects of number of lanes

With respect to the number of lanes, it was observed that 13.9% of pedestrians crossed on red at two-lane, 19% at three-lane and none at four-lane crosswalks. These differences were statistically significant (p = 0.00). Pedestrians were also found to decrease their walking speed when the number of lanes to be crossed decreased (p = 0.00). The average observed walking speed was 0.89 m/s on two-lane, 1.25 m/s on three-lane, and 1.37 m/s on four-lane crosswalks.

3.10. Other findings

Comparing the walking speed of pedestrians crossing on red and those crossing on green, it was observed that pedestrians who crossed on red walked faster (1.34 m/s on the average) than those crossing on green (1.16 m/s on the average). This difference was assessed using an unpaired single-sided t-test and was found to be statistically significant (p = 0.03).

Moreover, comparing those who crossed the road while carrying loads with other pedestrians, it was noticed that pedestrians carrying a load (63 or 8.9% out of 708 pedestrians) are more likely to walk slower than others with an average of 1.19 m/s for the former and 1.43 m/s for the latter. This difference was statistically significant using an unpaired single-sided t-test (p = 0.00).

Another finding was that only 207 (or 29.3%) out of the 708 pedestrians observed looked around before crossing the road. The majority of those observed, especially those who crossed on green, did not feel it was necessary to look around before crossing. Moreover and out of the 708 observed pedestrians, 569 (or 80.4%) were observed crossing the road within the pedestrian crossing designated area with the remaining 139 (19.6%) crossing fully or partly outside the designated area. As anticipated, pedestrians crossing inside the designated area tended to walk slower (average being 1.18 m/s) than those who crossed outside the designated area (average being 1.37 m/s). This difference was found to be statistically significant using an unpaired single-sided t-test with $p = 0.00$.

Finally and taking the behavior of using mobile phones while crossing into consideration, it was observed that pedestrians who crossed outside the designated area were more likely (13.67%) to use their mobile phone while crossing than those who were walking inside the area (6.33%). This difference was found to be statistically significant using an unpaired single-sided t-test with $p = 0.01$.

4. Discussion and conclusions

Analyzing pedestrian crossing behavior is important for ensuring their safety while crossing the roads. This study reviewed and examined the behaviors of a sample of 706 pedestrians crossing the road at 10 different signalized crosswalks in the city of Sharjah in UAE. The gender of pedestrians showed an effect on crossing speed where males take greater risk, on average, than females in crossing on red and walk, on average, faster than females. With respect to age, younger pedestrians were found to take greater risks than older pedestrians in terms of crossing on red and in terms of showing a more distractive behavior visualized in using the mobile phone while crossing the road.

The behavior of looking right and left before crossing was found to be affected by day of week where significantly more pedestrians were observed looking around before crossing on weekdays than on weekends when more people cross the road on average. One likely reason for this difference is that pedestrians feel that looking around before crossing the road is less necessary on weekends when more pedestrians cross the road than on weekdays. In other words, the greater number of pedestrians on weekends gives people crossing a greater sense of safety and, therefore, they feel looking around before crossing is less necessary than on weekdays. This possible reason goes along Le Bon's theory on group behavior. Le Bon (1895) said that when individuals join crowds, they become "deindividualized" where they take the behavior of the crowd. So when pedestrians feel that the group started crossing the street, they subconsciously join them with less time spent on perceiving the environment and deciding to cross.

Moreover, pedestrians crossing with larger groups walked at slower speeds, on the average, than those crossing with smaller groups or alone. One likely reason for that is the relying of pedestrians on groups to guide them while taking less time to perceive the environment (i.e., looking at the pedestrian light and for incoming traffic). This possible reason also goes along Le Bon's theory. So when pedestrians feel that a bigger crowd starts crossing the road, they subconsciously join them with less time spent on perceiving the environment and deciding to cross.

Results of the current study also showed that pedestrians tend to walk significantly more often outside designated areas and to cross on red more often at crosswalks with shorter green times and longer red times. It seems that longer red pedestrian lights lead to less willingness among pedestrians to wait longer times at crosswalks and, consequently, cross more often on red or outside the designated area (especially when they start crossing at the middle or towards the end of the green time) in order not to wait for the next green pedestrian light. Similar results were found in a study by Jain et al. (2014). The authors stated that red lights that were more than 40 s long made pedestrians look for alternatives to waiting at red, including crossing the road on red. Besides this, results of the current study indicated that pedestrians tend to walk significantly faster at crosswalks with shorter green times than at those with longer times. All such results clearly indicate that cycle time and red-green distribution affect the behavior of pedestrians and can lead them to disobey traffic rules. Such findings suggest that traffic authorities should keep red pedestrian lights duration to a minimum. If such an alternative is not feasible for any reason, then they should seek alternatives to signalized crosswalks like pedestrian bridges or underpasses.

With respect to the number of lanes, it was observed that some pedestrians are prepared to cross on red at two- and three-lane crosswalks but not at four-lane crosswalks. They were also found to walk faster as roads become wider. A likely reason for these differences is that wider roads make pedestrians feel more vulnerable and less safe if they walk slowly or cross on red, unlike narrower roads. Results indicate that the latter type of roads gives pedestrians greater, often false, confidence in their ability to cross safely, even on red pedestrian lights.

Results of the current study also indicated that pedestrians who crossed on red walked faster than those crossing on green. Such an outcome is anticipated as those crossing on green are more likely to walk at a normal pace, while those crossing on red are more likely to walk faster to avoid being hit by incoming motor vehicles. Moreover and as expected, pedestrians carrying loads and those walking with children are more likely to walk slower than others. It is therefore recommended that road designers take provisions (such as longer green duration) to ensure the safe crossing of pedestrians at signalized crosswalks where more children are likely to cross, like at schools and nurseries, and at those where more pedestrians are expected to carry loads, like in front of shopping centers.

One worrying and widespread behavior observed in the current study was that the majority (70.7%) of pedestrians observed did not look around before crossing the road. Another observed behavior by a small, but significant, proportion (or 19.6%) of pedestrians observed crossed the road fully or partially outside the designated area. Besides this, those crossing outside the designated area tended to walk faster than those who crossed outside the designated area. One likely reason for this difference in walking speed is that pedestrians walking outside the designated pedestrian area start crossing at the middle of the green time and feel the need to rush so that they can complete crossing before the light turns red.

One last finding was that almost 8% of pedestrians were observed crossing while using their mobile phones. In relation to this behavior, it was also observed that pedestrians who crossed outside the designated area were significantly more likely to use their mobile phone while crossing than those who were walking inside the area. This reflects that some pedestrians who are willing to take greater risk than others while crossing the road show this willingness in more than one way (i.e., using the phone and walking outside designated area in this case).

The Health Belief Model (HBM) can help in understanding the motivation of pedestrians in engaging in risky behaviors found in the current study (i.e., crossing on red, not looking around while crossing, and distracted crossing) and finding solutions that would minimize those risky behaviors. This model is a widely used behavioral change model that focuses on cognitive determinants of behavior. It states that people, in this case pedestrians, conduct an internal cognitive assessment to find out the net pros and costs of changing a certain behavior, and then decide whether or not to change this behavior (Carpenter, 2010). The original HBM has two components, threat perception and behavioral evaluation (Janz & Becker, 1984; Rosenstock et al., 1994). Then the model was extended by Sheeran and Abraham (1996) to include two more components, cues to action and health motivation. This model is occasionally used in understanding road users' behavior and proposing incentives to change them, like rear seatbelt use (Mehri et al., 2011) and bicycle helmet use (Lajunen & Räsänen, 2004).

To improve internal cognitive assessment and threat perception among pedestrians while crossing the road based on HBM, it is vital to improve their situation awareness. This can be done through rigorous public awareness campaigns to promote safe pedestrian road crossing behaviors and discouraging unsafe crossing behaviors using TV advertisements, social media, radio channels, newspapers, school curriculum, etc. Also, tightening road crossing laws, increasing police presence at pedestrian crosswalks, and increasing penalties would serve as vital incentives.

References

Alhajyaseen, W. K. M., & Iryo-Asano, M. (2017). Studying critical pedestrian behavioral changes for the safety assessment at signalized crosswalks. *Safety Science, 91*, 351–360.

Carpenter, C. J. (2010). A meta-analysis of the effectiveness of health belief model variables in predicting behavior. *Health Communication, 25*(8), 661–669.

Choi, J., Tay, R., Kim, S., & Jeong, S. (2019). Behaviors of older pedestrians at crosswalks in South Korea. *Accident Analysis & Prevention, 127*, 231–235.

Federal Competitiveness and Statistics Authority – United Arab Emirates [FCSA-UAE]. (2017). UAE in numbers. Abu Dhabi, UAE: FCSA.

Goh, B. H., Subramaniam, K., Wai, Y. T., & Mohamed, A. A. (2012). Pedestrian crossing speed: The case of Malaysia. *International Journal for Traffic and Transport Engineering, 2*, 323–332.

Jain, A., Gupta, A., & Rastogi, R. (2014). Pedestrian crossing behavior analysis at intersections. *International Journal for Traffic and Transport Engineering, 4*, 103–116.

Janz, N. K., & Becker, M. H. (1984). The health belief model: A decade later. *Health Education Quarterly, 11*, 1–47.

Koh, P. P., Wong, Y. D., & Chandrasekar, P. (2014). Safety evaluation of pedestrian behavior and violations at signalized pedestrian crossings. *Safety Science, 70*, 143–152.

Lajunen, T., & Räsänen, M. (2004). Can social psychological models be used to promote bicycle helmet use among teenagers? A comparison of the health belief model, theory of planned behavior and the locus of control. *Journal of Safety Research, 35*, 115–123.

Le Bon, G. (1895) The crowd: A study of the popular mind (Nye, R.A., Trans.). London: Ernest Benn (translation published 1947).

Mehri, A., Nadrian, H., Morowatisharifabad, M. A., & Akolechi, M. (2011). Determinants of seat belt use among drivers in Sabzevar, Iran: A comparison of theory of planned behavior and health belief model. *Traffic Injury Prevention, 12*, 104–109.

Mwakalonge, J., Siuhi, S., & White, J. (2015). Distracted walking: Examining the extent to pedestrian safety problems. *Journal of Traffic and Transportation Engineering (English Edition), 2*, 327–337.

Nordfjærn, T., & Şimşekoğlu, Ö. (2013). The role of cultural factors and attitudes for pedestrian behavior in an urban Turkish sample. *Transportation Research Part F: Traffic Psychology and Behavior, 21*, 181–193.

Papadimitriou, E., Theofilatos, A., & Yannis, G. (2013). Patterns of pedestrian attitudes, perceptions and behavior in Europe. *Safety Science, 53*, 114–122.

Rosenstock, I. M., Strecher, V. J. and Becker, M. H. (1994). The health belief model and hiv risk behavior change. In DiClemente, R. J. and Peterson, J. L. (Eds.), Preventing AIDS, theories and methods of behavioral interventions. Springer, 5–24.

Sheeran, P., & Abraham, C. (1996). The health belief model. In M. Conner & P. Norman (Eds.), Predicting health behavior. Buckingham: Open University Press, 23–62.

Tezcan, H. O., Elmorssy, M., & Aksoy, G. (2019). Pedestrian crossing behavior at midblock crosswalks. *Journal of Safety Research, 71*, 49–57.

Thompson, L. L., Rivara, F. P., Ayyagari, R. C., & Ebel, B. E. (2013). Impact of social and technological distraction on pedestrian crossing behavior: An observational study. *Injury Prevention, 19*, 232–237.

World Health Organization. (2018). Global status report on road safety 2018. Geneva, Switzerland: World Health Organization.

Zegeer, C. V., & Bushell, M. (2012). Pedestrian crash trends and potential countermeasures from around the world. *Accident Analysis & Prevention, 44*(1), 3–11.

Dr. Salaheddine Bendak holds BSc and MSc in Industrial Engineering from the Middle East Technical University (Turkey), Graduate Diploma of Education from Monash University (Australia), graduate certificate in the Global Burden of Injuries from Karolinska Institute (Sweden) and a PhD in Ergonomics from La Trobe University (Australia). He worked in five different countries as an industrial engineer, occupational/transportation safety consultant and university professor. He is currently working as a professor at the Department of Industrial Engineering, Halic University, Turkey. His research interests include transportation safety, occupational safety, applied ergonomics, safety culture and organisational behaviour.

Asayel M. Alnaqbi holds a BSc in Industrial Engineering and Engineering Management from the University of Sharjah in 2019. Currently working in various public sector organisations in Sharjah, United Arab Emirates.

Muna Y. Alzarooni holds a BSc in Industrial Engineering and Engineering Management from the University of Sharjah in 2019. Currently working in various public sector organisations in Sharjah, United Arab Emirates.

Sara M. Aljanaahi holds a BSc in Industrial Engineering and Engineering Management from the University of Sharjah in 2019. Currently working in various public sector organisations in Sharjah, United Arab Emirates.

Shaikha J. Alsuwaidi holds a BSc in Industrial Engineering and Engineering Management from the University of Sharjah in 2019. Currently working in various public sector organisations in Sharjah, United Arab Emirates.

Journal of Safety Research 76 (2021) 44–55

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Highway safety assessment and improvement through crash prediction by injury severity and vehicle damage using Multivariate Poisson-Lognormal model and Joint Negative Binomial-Generalized Ordered Probit Fractional Split model

Kai Wang [a,*], Tanmoy Bhowmik [b], Shanshan Zhao [a], Naveen Eluru [b], Eric Jackson [a]

[a] *Connecticut Transportation Safety Research Center, Connecticut Transportation Institute, University of Connecticut, United States*
[b] *Department of Civil, Environmental & Construction Engineering, University of Central Florida, United States*

A R T I C L E I N F O

Article history:
Received 31 January 2020
Received in revised form 11 August 2020
Accepted 17 November 2020
Available online 14 December 2020

Keywords:
Crash prediction by severity
Crash severity surrogate
Unobserved heterogeneity
Multivariate model
Joint fractional split model

A B S T R A C T

Introduction: Predicting crash counts by severity plays a dominant role in identifying roadway sites that experience overrepresented crashes, or an increase in the potential for crashes with higher severity levels. Valid and reliable methodologies for predicting highway accidents by severity are necessary in assessing contributing factors to severe highway crashes, and assisting the practitioners in allocating safety improvement resources. *Methods:* This paper uses urban and suburban intersection data in Connecticut, along with two sophisticated modeling approaches, *i.e.* a Multivariate Poisson-Lognormal (MVPLN) model and a Joint Negative Binomial-Generalized Ordered Probit Fractional Split (NB-GOPFS) model to assess the methodological rationality and accuracy by accommodating for the unobserved factors in predicting crash counts by severity level. Furthermore, crash prediction models based on vehicle damage level are estimated using the same two methodologies to supplement the injury severity in estimating crashes by severity when the sample mean of severe injury crashes (*e.g.*, fatal crashes) is very low. *Results:* The model estimation results highlight the presence of correlations of crash counts among severity levels, as well as the crash counts in total and crash proportions by different severity levels. A comparison of results indicates that injury severity and vehicle damage are highly consistent. *Conclusions:* Crash severity counts are significantly correlated and should be accommodated in crash prediction models. *Practical application:* The findings of this research could help select sound and reliable methodologies for predicting highway accidents by injury severity. When crash data samples have challenges associated with the low observed sampling rates for severe injury crashes, this research also confirmed that vehicle damage can be appropriate as an alternative to injury severity in crash prediction by severity.

1. Introduction

1.1. Motivation

Each year there are over 38,000 motor-vehicle crashes related fatalities in the United States of America, and traffic collisions are one of the most significant causes of untimely death (NHTSA, 2016). Traffic safety is a top priority for both Federal and State transportation agencies and there is still a critical need for effective strategies to reduce crashes and improve highway safety.

Crash prediction models are one of the most effective approaches to help identify roadway locations with overrepresented crashes or the potential for crashes in the future. These predictive model results can then be used to implement countermeasures to improve highway safety. Therefore, selecting an appropriate and effective crash prediction model is critical when trying to identify roadway sites to prioritize for safety improvement. The use of inaccurate or invalid modeling approaches and assumptions might result in biased crash prediction results and thus lead to the inefficient use of safety improvement resources and reduce the effectiveness of the safety management process. Given the limited safety improvement resources available, sites that experience overrepresented high severity crashes should be our top priority. The development of reliable crash prediction methodologies, based on crash severity,

* Corresponding author.
E-mail addresses: kai.wang@uconn.edu (K. Wang), tanmoy78@Knights.ucf.edu (T. Bhowmik), shanshan.h.zhao@uconn.edu (S. Zhao), naveen.eluru@ucf.edu (N. Eluru), eric.d.jackson@uconn.edu (E. Jackson).

https://doi.org/10.1016/j.jsr.2020.11.005
0022-4375/

is imperative for helping to identify hazardous roadway locations and crash contributing factors. This allows for the efficient allocation of highway safety improvement strategies to assist in preventing crashes from occurring in the future.

The first edition of the Highway Safety Manual (HSM, 2010) introduces the crash prediction models for total crashes, which are then multiplied by several constant severity proportions to predict the crashes by different severity levels. This approach might not be feasible as crash severity distributions may vary across sites due to potential variations related to roadway geometric, traffic, and environmental characteristics (Ma & Kockelman, 2006; Ma, Kockelman, & Damien, 2008; Wang, Ivan, Burnicki, & Mamun, 2017; Wang, Zhao, & Jackson, 2018, 2019; Wang, Ivan, Ravishanker, & Jackson, 2017). Therefore, crash prediction models by severity have been widely investigated to improve the prediction performance of crash counts by different severity levels (Abdel-Aty & Radwan, 2000; Dixon et al., 2015; Liu & Sharma, 2018; Lord & Persaud, 2000; Oh, Washington, & Choi, 2004; Russo, Busiello, & Dell, 2016; Tarko, Inerowicz, Ramos, & Li, 2008; Ulfarsson & Shankar, 2002). In general, two types of methodological frameworks of crash prediction models have been implemented by researchers to achieve a better crash severity count prediction. The first alternative is to estimate crash counts by different severity levels directly. The second alternative, which is usually referred to as the two-stage model, is first to estimate crash counts in total, followed by estimating crash severity distributions, and then combine the latter with the former for crash count prediction by severity.

Estimating crash prediction models by severity might be challenging due to the small sample size and low sample mean (Anarkooli, Persaud, Hosseinpour, & Saleem, 2019), especially for the fatal and severe injury crashes. This alternative creates an issue for identifying locations with overrepresented severe crashes when the safety improvement resources are limited. To this end, one objective indicator of crash consequences – the extent of vehicle damage based on the destruction/deformation of the vehicle involved in the crash might be used – to represent the crash consequence as a supplement of injury severity. The rationality of modeling crashes by vehicle damage level is because: (a) the sample mean of crashes with severe vehicle damage levels is higher than the crashes with severe injuries; and (b) the vehicle damage is found to be positively correlated with the injury severity in multiple studies (Qin, Sultana, Chitturi, & Noyce, 2013; Wang et al., 2015, 2019). For this reason, roadway locations experienced overrepresented crashes with severe vehicle damage levels have a very high potential to experience more severe injury crashes in the future.

1.2. Literature review

With regard to the methodologies that directly estimate crash counts by severity, the Poisson regression model has been initially used to model crashes by each severity level since the crash frequencies are non-negative integers (Lord & Mannering, 2010). The Poisson model has its implicit restriction – the variance of the data is assumed to be equal to the mean. This assumption might not always be valid as the variance of crash data usually is higher than the mean, which is also known as the over-dispersion (Washington, Karlaftis, & Mannering, 2011). To address this issue, the Univariate Poisson Lognormal regression and Negative Binomial regression models are then used to predict crash counts by severity (Mannering & Bhat, 2014; Washington et al., 2011). However, traditional Univariate Poisson Lognormal and Negative Binomial models assume crash counts by crash severity to be independent. However, this might not be true due to the presence of shared unobserved factors across different severity levels for each observational record. Modeling crash severity counts together without accounting for their correlations might yield biased parameter estimates, and reduce model prediction accuracy (Ma & Kockelman, 2006; Ma et al., 2008; Mannering & Bhat, 2014; Mannering, Shankar, & Bhat, 2016; Wang et al., 2017, 2018).

To address correlations among crash counts across different severity levels, a large number of methodologies have been implemented to estimate crash counts by severity jointly. These include, but are not limited to, Simultaneous Equations Model (Ye, Pendyala, Shankar, & Konduri, 2013); Multivariate Generalized Poisson Model (Chiou & Fu, 2013, 2015; Chiou, Fu, & Chih-Wei, 2014); Joint-Probability Model (Pei, Wong, & Sze, 2011); and Artificial Neural Network (Zeng, Huang, Pei, & Wong, 2016). Recently, the Multivariate regression models have been extensively applied to simultaneously estimate crash counts by severity by accounting for the correlations between crashes among different crash severity levels. Multivariate models have been verified to be superior to the Univariate models in terms of the parameter estimation and crash prediction accuracy. For instance, Ma et al. (2008) applied a Multivariate Poisson-Lognormal model to estimate the crash counts by severity, and they found the crash counts highly correlated among different severity levels. Park and Lord (2007) applied a Multivariate Poisson-Lognormal model to jointly estimate the crash frequencies by severity using the California data. The study implied that the crash frequencies are highly correlated among severity levels, and the Multivariate model obtains more accurate parameter estimates. Wang, Ivan, Ravishanker, and Jackson (2017) used a Multivariate Lognormal approach to estimate crash count models for rural two-lane undivided highways, and the results were compared to the Univariate models. The study verified that the Multivariate Lognormal model provides unbiased parameter estimates and significantly enhances the prediction accuracy. A similar study was conducted by Wang, Zhao, and Jackson (2018) for freeway crash prediction, and the results highlighted that freeway crashes significantly correlated among different levels of crash severity. Anastasopoulos, Shankar, Haddock, and Mannering (2012) used both Multivariate Tobit and Multivariate Negative Binomial models to predict crash rate by severity on multilane divided highways in Washington State. The study found that the prediction accuracy between the two approaches are very close, and both methods outperform the univariate models.

Furthermore, the Multivariate models have also been extended by researchers for particular perspectives. For instance, to account for the issues of excess zero, unobserved heterogeneity and spatial–temporal correlation in crash data, methodologies (including but are not limited to) Multivariate Random-Parameter Zero-Inflated model (Dong, Clarke, Yan, Khattak, & Huang, 2014), Multivariate Poisson Lognormal Spatial model (Aguero-Valverde, 2013; Barua, El-Basyouny, & Islam, 2014) , Multivariate Spatial-Temporal Bayesian model (Liu & Sharma, 2018), Multivariate Poisson Lognormal Conditional-Autoregressive model (Wang & Kockelman, 2013; Xie, Ozbay, & Yang, 2019) and Multivariate Random Parameter Spatial Poisson Lognormal model (Barua, El-Basyouny, & Islam, 2016) were then used to estimate the crash counts by severity. Lord and Mannering (Lord & Mannering, 2010) provided comprehensive guidance on model selection and assessment in crash count prediction.

Now on to the second alternative approach described above. Qin, Wang, and Cutler (2013) used a Negative Binomial model and a Multinomial Logit model to predict total truck crashes and crash counts by each severity level. Chiou and Fu (2013, 2015)) and Chiou et al. (2014) examined the use of a Multinomial model and a Generalized Poisson model to predict crash frequencies by severity. Anarkooli et al. (2019) applied a Negative Binomial model and a Generalized Ordered Probit model to estimate crashes by severity on horizontal curves. Geedipally, Bonneson, Pratt, and Lord (2013) used a Multinomial Logit model to estimate the severity distributions for freeway segments and interchanges. Wang, Quddus, and Ison (2011) applied a Bayesian spatial model and a mixed logit model to estimate crashes by severity for the major roads in England. Savolainen, Mannering, Lord, and Quddus (2011) provided comprehensive guidance on model selection and assessment of prediction of crash severity distributions.

However, all of these studies modeled the total crash counts and crash severity distributions separately and independently, which might be inappropriate due to the common observed and unobserved factors that affect both crash counts and crash severity distributions. Yasmin, Eluru, Lee, and Abdel-Aty (2016) introduced a new modeling framework – The Joint Negative Binomial-Ordered Probit Fractional Split (NB-OPFS) model to estimate the total crashes and crash severity distributions simultaneously. In their method, a Negative Binomial component was employed for estimating crashes in total, and an Ordered Probit fractional split component was employed for estimating crash proportions by severity. Unlike previous studies, their modeling framework jointly estimates the Negative Binomial component and the Ordered Probit component, by accounting for the unobserved heterogeneity across and within the crash count and crash severity proportion modeling components. Further, by implementing the Ordered Probit framework, the method also accounts for the ordinal nature of crash severity in the crash proportion estimation.

The authors then further extended their methodology (Yasmin & Eluru, 2018) to a Joint Negative Binomial-Generalized Ordered Logit Fractional Split modeling framework to estimate crash counts by severity at a zonal level for Florida State. This method allowed the correlation between total crash counts and crash severity proportions to vary across zones. The study highlighted the superiority of the joint model framework in terms of the prediction accuracy compared to the independent model framework. Bhowmik, Yasmin, and Eluru (2019) applied a Panel Mixed Generalized Ordered Probit Fractional Split model to examine the contributing factors to vehicle operating speed. The study found that roadway related characteristics significantly affect the vehicular speed, and the proposed model framework performs adequately for the speed prediction.

1.3. Problem statement, study objectives and contributions

Although different methodologies have been applied in predicting crashes by severity, multiple issues are still existent and need to be addressed. For instance:

1. Most of the previous studies focused on implementing one of the two options for crash prediction by severity, *i.e.* either predicting crash severity counts simultaneously or predicting total crash counts and crash severity proportions together. There is a lack of study in assessing and comparing these two options in highway safety research, which can offer insights on the pros and cons of each method, and shed light on method selection under different data conditions and research needs.
2. Although previous research has verified that the low sample mean of severe crashes (*e.g.*, fatal crashes) leads to difficulties in crash prediction by severity level, limited research provided effective alternatives to address this issue. The shortage of crash prediction capability for severe crashes creates troubles to practitioners for identifying target roadway locations, when the highway safety improvement resources are limited.

Accordingly, three major objectives and contributions are addressed and made by this study respectively. They are:

1. Assess and identify the most reliable methodology in predicting crashes by severity, using and extending the two advanced statistical methodologies, *i.e.* the Multivariate Poisson-Lognormal (MVPLN) model and the Joint Negative Binomial-Generalized Ordered Probit Fractional Split (NB-GOPFS) model.
2. Identify and interpret the contributing factors to severe crashes.
3. Evaluate the rationality of using the vehicle damage as an alternative to injury severity in crash prediction models by severity level, to provide practitioners with capabilities of effectively allocating safety improvement resources when the low sample mean leads to difficulties in predicting severe crashes.

The remaining parts of this paper are as follows: the second section describes the two methodological frameworks and the model estimation methods; the third section describes the data used in model estimation; and the fourth section discusses the model estimation results. Model comparisons are provided in the fifth section and conclusions are discussed in the final section.

2. Methodologies

2.1. Framework for Multivariate Poisson-Lognormal (MVPLN) model

The first method used to estimate crash counts by severity is the MVPLN model. Assume $y_i = (Y_{1i}, Y_{2i} \cdots, Y_{Ji})'$ for $i = 1, 2, \cdots, N$ be a *J*-dimensional vector (*i.e. J* crash severity levels) of crash counts across all *N* sites. In the MVPLN model, we assume the crash counts are correlated among all severity levels. The MVPLN model can be derived as (Serhiyenko, Mamun, Ivan, & Ravishanker, 2016):

$$Y_{ji}|\lambda_{ji}\ Poisson(\lambda_{ji}) \tag{1}$$

where λ_{ji} is the mean of Poisson distribution, which is estimated as:

$$ln(\lambda_{ji}) = Offset + \boldsymbol{\beta}_j \boldsymbol{x_{ji}} + \varepsilon_{ji} \tag{2}$$

where *Offset* is the log exposure for total observation days in the data set for intersection models (*i.e.*, in this study, the offset for both sign-controlled and signalized intersections is log(365*5) = 7.51). $\boldsymbol{x}_{ji}$ is a vector of independent variables and $\boldsymbol{\beta}_j$ is a vector of coefficients to be estimated. ε_{ji} is a random term. Assume a vector of the random term $\boldsymbol{\varepsilon}_i = (\varepsilon_{1i}, \varepsilon_{2i} \cdots, \varepsilon_{Ji})'$ at site *i* follows a *J*-dimensional normal distribution, *i.e.*

$$\boldsymbol{\varepsilon_i}\ Normal(0, \Sigma) \tag{3}$$

where **0** is a *J*-dimensional zero vector, Σ is a J *J variance–covariance matrix and let's define $\Sigma = (\sigma_{rs})_{1 \le r < s \le J}$. Then the mean, variance and covariance of the crash counts by each severity level at site *i* can be derived as (Serhiyenko et al., 2016):

$$Mean = E[Y_{ji}] = exp(Offset + \boldsymbol{\beta}_j \boldsymbol{x}_{ji}) exp(\frac{\sigma_{jj}}{2}) \tag{4}$$

$$\begin{aligned} Variance = Var[Y_{ji}] = exp\left(Offset + \boldsymbol{\beta}_j \boldsymbol{x}_{ji}\right) exp\left(\frac{\sigma_{jj}}{2}\right) \\ + exp\left(2\left(Offset + \boldsymbol{\beta}_j \boldsymbol{x}_{ji}\right)(exp^2(\sigma_{jj}) - exp(\sigma_{jj})\right) \end{aligned} \tag{5}$$

$$\begin{aligned} Covariance = Cov[Y_{ri}, Y_{si}] \\ = exp(Offset + \boldsymbol{\beta}_j \boldsymbol{x}_{ji}) exp(\frac{\sigma_{rr}}{2}) exp(\frac{\sigma_{ss}}{2})(exp(\sigma_{rs}) - 1) \end{aligned} \tag{6}$$

The correlations of crash counts between r^{th} and s^{th} crash severity can be accommodated by the covariance term in equation (Wang et al., 2017) through the σ_{rs}, which is the off-diagonal entry of the J *J variance–covariance matrix. A positive σ_{rs} represents a positive correlation of crash counts between r^{th} and s^{th} crash severity, and a negative σ_{rs} represents a negative correlation of crash counts between r^{th} and s^{th} crash severity. The σ_{rs} can be further derived as:

$$\sigma_{rs} = \rho_{rs}\sqrt{\sigma_{ss} * \sigma_{rr}} \tag{7}$$

where σ_{ss}andσ_{rr} are the diagonal entries of the J *J variance–covariance matrix, and ρ_{rs} is a traditional correlation coefficient to be estimated which is between −1 and 1. The probability distribution of the given total crash counts $\boldsymbol{y}_i$ can be written as (Serhiyenko et al., 2016):

$$g(\boldsymbol{y}_i|\boldsymbol{\beta}_j\boldsymbol{x}_{ji},\Sigma) = \int\cdots\int f_{Normal,J}(\boldsymbol{\varepsilon}_i|0,\Sigma)\prod_{j=1}^{J} f_{Poisson}\left(y_{ij}\middle|\varepsilon_{ji},\boldsymbol{\beta}_j\boldsymbol{x}_{ji}\right)d\boldsymbol{\varepsilon}_i \tag{8}$$

where $f_{Normal,J}$ is a J-dimensional normal distribution function, and $f_{Poisson}$ is a Poisson distribution. As noted from previous studies (Serhiyenko et al., 2016; Wang et al., 2017), the probability distribution function shown in equation 8 has no closed algebraic solution, and hence the Bayesian framework is used to estimate the coefficients in the MVPLN model. First assume every $\boldsymbol{\beta}_j$ in equation 2 follows a prior normal distribution as *Normal* $(0, 6^2)$ and every $\boldsymbol{\Sigma}^{-1}$ in equation 3 follows a prior Wishart distribution as *Wishart* (c, ϖ), where 6^2, c and ϖ are all hyperparameters for priors. We used the default hyperparameter specifications in R-INLA (2020) for both the Normal prior(*i.e.*β_j $Normal\left(0,10^3\right)$) and the Wishart prior with $c = 7$ (*i.e.* $2J + 1$) degrees of freedom and an identify matric as the precision matrix ϖ. The posterior distributions of the coefficients are estimated using the Bayesian inference (Serhiyenko et al., 2016).

The Markov Chain Monte Carlo (MCMC) (Ma & Kockelman, 2006; Ma et al., 2008) simulation, which uses the Gibbs sampler and Metropolis-Hasting (M-H) approach, is usually applied to carry out the Bayesian inference on model estimation. However, studies have noticed that the MCMC simulation approach is extremely computationally challenging and time-consuming, especially for a large data sample (Mannering & Bhat, 2014). To address this issue and simplify the model estimation procedure, we applied the Integrated Nested Laplace Approximation (INLA) approach proposed by Rue, Martino, and Chopin (2009) to carry out the Bayesian inference of the MVPLN model estimation in this study. The INLA approach doesn't rely on the MCMC and it numerically approximates the posterior distributions of parameters. It has been verified to be able to significantly reduce the running time compared to the MCMC approach by multiple studies (Serhiyenko et al., 2016; Wang et al., 2017, 2018). The R-INLA (2020) package was used to run the MVPLN models. The detailed discussions of the INLA approach and model estimation procedures are referenced in several previous studies (Rue et al., 2009; Serhiyenko et al., 2016).

2.2. Framework for Joint Negative Binomial-Generalized Ordered Probit Fractional Split (NB-GOPFS) Model

In the NB-GOPFS model, the total crash counts and crash proportions by each severity are jointly estimated, by accounting for the correlations between total crashes and crash severity proportions. Therefore, two model components are included in the NB-GOPFS method, where a count model (*i.e.*, a Negative Binomial framework is used in this study) is used to estimate the total crash counts, and a fractional split model (*i.e.*, a Generalized Ordered Probit Fractional Split framework) is used to estimate the crash proportions by each severity level. Similar to the MVPLN model framework, assume i ($i = 1, 2, 3 \ldots N$) to be the index for the roadway site, and j ($j = 1, 2, 3 \ldots J$) to be the index for the injury severity category. The total crash counts y_i at site i can be estimated using a NB framework, which is derived as:

$$Prob[y_i|\mu_i] = p(y_i) = \frac{\Gamma[(\sigma)+y_i]}{\Gamma(\sigma)y_i!}\left[\frac{\sigma}{(\sigma)+\mu_i}\right]^{\sigma}\left[\frac{\mu_i}{(\sigma)+\mu_i}\right]^{y_i} \tag{9}$$

where Γ is a gamma function; σ is the inverse overdispersion parameter in the NB model, and μ_i is the expected crash counts at site i, which can be written as:

$$ln(\mu_i) = Offset + (\boldsymbol{\beta} + \boldsymbol{\zeta_i})\boldsymbol{x}_i + \varepsilon_i + \eta_i \tag{10}$$

where $\boldsymbol{x}_i$ is a vector of independent variables associated with site i and $\boldsymbol{\beta}$ (not including a constant) is a vector of coefficients to be estimated. $\boldsymbol{\zeta_i}$ (which follows a standard normal distribution: $\boldsymbol{\zeta_i}$ $N(0,\boldsymbol{\pi}^2)$) is a vector of estimated coefficients which accounts for the unobserved heterogeneity in crash count estimation at site i. exp(ε_i) is a random term that follows a gamma distribution with mean 1 and variance σ. η_i is a random factor that accommodates the correlations between total crash counts and crash severity proportions at site i, due to the common unobserved factors.

Considering the ordinal nature of crash severity, the estimation of proportions by each crash severity level is carried out by a Generalized Ordered Probit Fractional Split (GOPFS) framework. Let's define the p_{ji} be the actual proportion of crash severity j at site i, which is assumed to be associated with a latent variable $p_i{}^*$. The latent variable can be specified as (Yasmin & Eluru, 2018):

$$p_i{}^* = (\gamma + \boldsymbol{\rho_i})\boldsymbol{z}_i + \delta_i + \eta_i, \;\; p_{ji} = j \; if \, \epsilon_{i,j-1} < p_i{}^* < \epsilon_{i,j} \tag{11}$$

This latent propensity $p_i{}^*$ is mapped to the actual severity proportion categories p_{ji} by the ϵ thresholds ($\epsilon_0 = -\infty$ and $\epsilon_j = \infty$). $\boldsymbol{z}_i$ is a vector of attributes that influences the propensity associated with crash severity proportions. $\boldsymbol{\gamma}$ is a corresponding vector of mean effects, and $\boldsymbol{\rho_i}$ is a vector of unobserved factors on severity proportion propensity for site i and its associated characteristics assumed to be a realization from standard normal distribution: $\boldsymbol{\rho_i}$ $N(0,\boldsymbol{k}^2)$. δ_i is an idiosyncratic random error term assumed to be identically and independently standard normal distributed across observational unit i.η_i is a random factor that accommodates the correlations between total crash counts and crash severity proportions at site i, due to the common unobserved factors.

The GOPFS model relaxes the constant thresholds across observations to provide a flexible form of the OPFS model. The basic idea of the GOPFS is to represent the threshold parameters as a linear function of exogenous variables to account for the heterogeneity. Thus, the thresholds are expressed as:

$$\epsilon_{i,j} = fn(s_{ij}) \tag{12}$$

where, s_{ij} is a set of exogenous variables (including a constant) associated with j^{th} threshold. Further, to ensure the accepted ordering of observed severity proportions $(-\infty < \epsilon_{i,1} < \epsilon_{i,2} < \cdots\cdots\cdots < \epsilon_{i,J-1} < +\infty)$, we use the following parametric form as employed by Eluru, Bhat, and Hensher (2008):

$$\epsilon_{i,j} = \epsilon_{i,j-1} + exp((\tau_j + \theta_{ji})s_{ij} + \eta_i \tag{13}$$

where, τ_j is a vector of parameters to be estimated. θ_{ji} is another vector of unobserved factors moderating the influence of attributes in s_{ij} on the severity proportions for analysis unit i and injury severity category j. It is noted from equation 11 that p_{ji} is the actual proportion of crash severity j, which is different from the traditional generalized ordered Probit model framework where the dependent variable is an indicator of crash severity level. In order to estimate the generalized order Probit framework with a continuous dependent variable, let's assume (Yasmin & Eluru, 2018)

$$E(p_{ji}|\boldsymbol{x}_i) = H_{ji}(\gamma, \epsilon), \;\; 0 \le H_{ji} \le 1, \sum_{j=1}^{J} H_{ji} = 1 \tag{14}$$

H_{ji} accounts for the ordered Probit probability (P_{ji}) form for the crash severity level *j*, and it is defined as:

$$P_{ji} = \phi\{\epsilon_{i,j} - [(\boldsymbol{\gamma} + \boldsymbol{\rho_i})\boldsymbol{z}_i + \delta_i + \eta_i]\} - \phi\{\epsilon_{i,j-1} - [(\boldsymbol{\gamma} + \boldsymbol{\rho_i})\boldsymbol{z}_i + \delta_i + \eta_i]\} \tag{15}$$

where ϕ is the cumulative standard normal distribution. It is noted from previous research (Yasmin & Eluru, 2018) that the correlations between total crash counts and crash proportions by severity may vary across sites. Therefore, we parameterize the correlation parameter in this study as follows:

$$\eta_i = \boldsymbol{\alpha c}_i \tag{16}$$

where, $\boldsymbol{c}_i$ is a vector of exogenous variables, $\boldsymbol{\alpha}$ is a vector of unknown parameters to be estimated (including a constant).

To jointly estimate the NB probability function (see equation 9) for total crash counts and the GOPFS probability function (see equation 15), let's define a structure Ω for all vectors (*i.e.*, $\zeta, \boldsymbol{\rho}, \boldsymbol{\theta}$ and $\boldsymbol{\alpha}$) that account for unobserved heterogeneity, either in NB or GOPFS model framework, and Ω $N(0, (\boldsymbol{\pi}^2, \boldsymbol{k}^2, \boldsymbol{m}^2, \boldsymbol{n}^2))$. The likelihood function of the Joint NB-GOPFS model can be written as:

$$L_i = \int_{\Omega} p(y_i) \times \prod_{j=1}^{J} P_{ji}^{\omega_i p_{ji}} d\Omega \tag{17}$$

where ω_i is a dummy indicator where $\omega_i = 1$ represents site *i* has at least one crash, otherwise $\omega_i = 0$. The log-likelihood function can then be written as:

$$LL = \sum_i Ln(L_i) \tag{18}$$

Overall, the parameters to be estimated in the Joint NB-GOPFS model are $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\sigma}$, ϵ, $\boldsymbol{\pi}$, $\boldsymbol{k}$, $\boldsymbol{m}$ and $\boldsymbol{n}$. The Quasi-Monte Carlo simulation approach based on the scrambled Halton sequence is applied to estimate the log-likelihood function, using the GAUSS Matrix Programming Software (Aptech, 2019). The detailed discussions of the Joint NB-GOPFS approach and model estimation procedures are referenced in several previous studies (Bhat, 2001; Bhowmik et al., 2019; Eluru et al., 2008; Yasmin & Eluru, 2013, 2018; Yasmin et al., 2016).

3. Data preparation

To estimate and compare the MVPLN and Joint NB-GOPFS models for crash prediction by severity, urban & suburban intersections were collected from the State of Connecticut and five-year crash data (2014–2018) were collected from the Connecticut Transportation Crash Data Repository (CTCDR) (2019) and assigned to the specific intersections.

A total of 895 intersections are sign-controlled and 1,178 are signalized. To obtain sufficient observations in each crash severity level, crash severity counts were aggregated into three categories (Wang et al., 2017, 2018):

1) K + A which combines fatal (K) and incapacitating injury (A) crashes;
2) B + C which combines non-incapacitating injury (B) and possible injury crashes (C), and
3) PDO which includes the property damage only (PDO) crashes.

As mentioned earlier, vehicle damage is used as another crash consequence indicator to supplement the crash injury severity in this study to further address the low sample mean issue in crash prediction models, especially for estimating models for crashes with severe injuries such as K and A crashes. According to the Model Minimum Uniform Crash Criteria (MMUCC) guideline (2017), vehicle damage was categorized into five levels and vehicle damage counts were aggregated into three categories in this study (Qin et al., 2013; Wang et al., 2015, 2019):

1) Severe Damage Crashes which contain all crashes with disabling (salvageable or total loss) damage;
2) Moderate Damage Crashes which contain all crashes with broken or missing parts damage, and
3) Minor Damage Crashes which combine crashes with minor/cosmetic damage and crashes with no damage.

To validate the assumption of using vehicle damage to supplement the injury severity to address the low sample mean issue in crash prediction models, Fig. 1 presents the scatter plots and Pearson correlation coefficients between the injury severity and vehicle damage for both sign-controlled and signalized intersections. The Pearson correlation coefficients illustrated that the injury severity and vehicle damage are highly correlated for both sign-controlled and signalized intersections. Furthermore, intersection traffic and geometric data were collected based on the urban & suburban arterials chapter in the Highway Safety Manual (HSM) (2010) with regard to both sign-controlled and signalized intersections. Table 1 summarizes the descriptive characteristics of the intersection and crash data used in this study.

4. Model estimation results

Tables 2 and 3 show the estimation results for the MVPLN and Joint NB-GOPFS models by different injury severity and vehicle damage levels for urban & suburban sign-controlled and signalized intersections, respectively. In each cell, the first value represents the estimated coefficient, followed by the *p*-value of the coefficient in parenthesis. "–" represents the coefficient is not statistically significant at the 10% significance level, and the results only included variables that are significant at least in one of the models. "NA" represents the variable is not applicable in the specific model.

4.1. Urban & suburban sign-controlled intersections

Table 2 presents the model estimation results for urban & suburban sign-controlled intersections. The upper part shows the estimated coefficients of the crash prediction models by injury severity level, and the lower part shows the estimated coefficients of the crash prediction models by vehicle damage level.

4.1.1. Model estimation for injury severity component

With regard to the MVPLN model by injury severity, the crash counts by three different severity levels (*i.e.* K + A, B + C and PDO crashes) are simultaneously estimated using a Poisson-Lognormal framework by accounting for their correlations due to the common unobserved factors. Both the major and minor road AADT are found to be statistically significant and are positively associated with all three levels of crash severity counts. Compared with 3-leg intersections, 4-leg intersections are associated with increased crash counts by all severity levels, which may be due to the fact that there are more conflicting points at 4-leg intersections. As expected, all-way stop-controlled intersections have experienced decreased crashes with severe injuries because right-of-way is separated for all approaches, and the vehicle speed is lower as all vehicles are ordered to stop first and then go at all-way stop-controlled intersections. If a driveway (such as a driveway for gasoline, park-

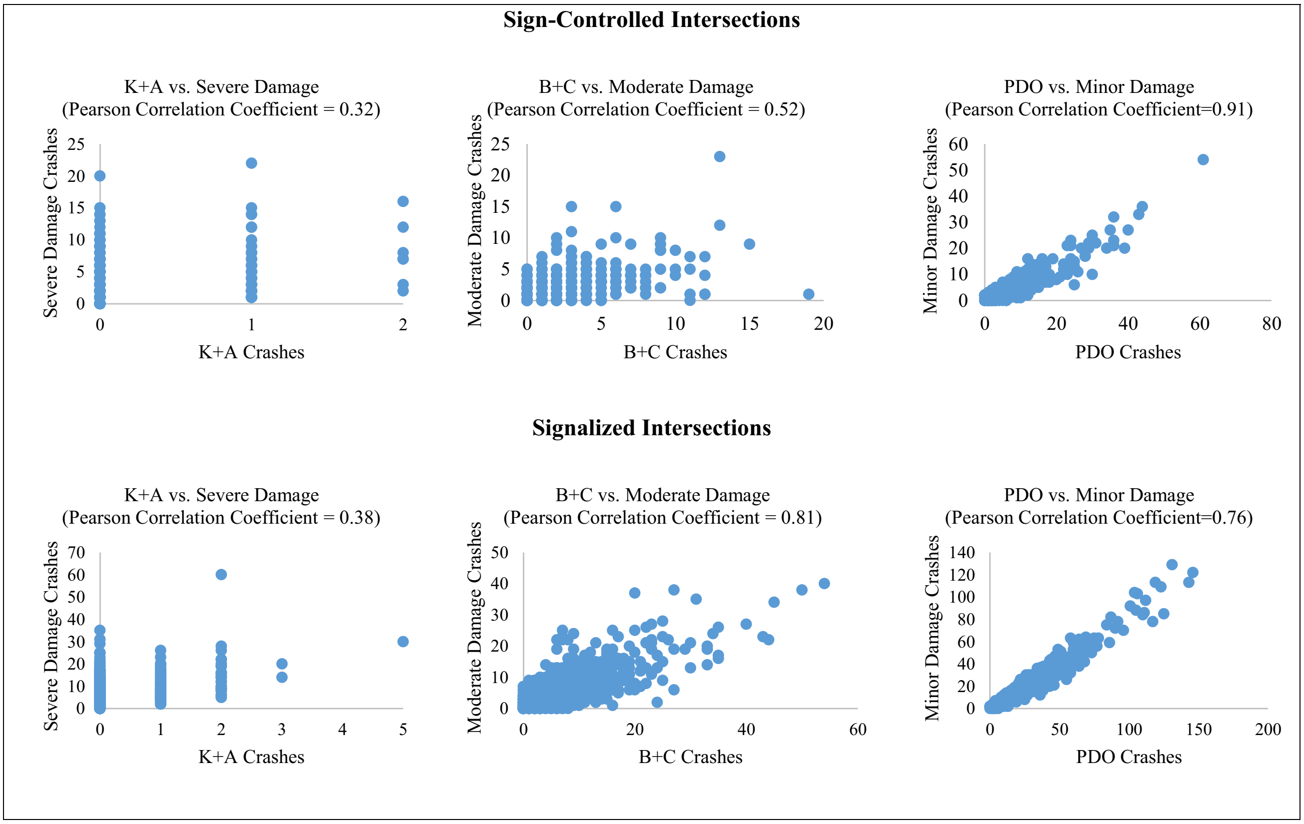


Fig. 1. Correlation between Injury Severity and Vehicle Damage.

ing, or commercial store) is present at the intersection, crash counts by all three severity levels are expected to decrease. This might be because drivers tend to drive more carefully at these intersections where vehicles may exit from the nearby driveway. Exclusive left-turn lanes are associated with decreased crash counts for all three severity levels, and exclusive right-turn lanes are associated with increased crashes with less severe injuries. These findings are consistent with the study conducted by Wang et al. (2017) that exclusive left-turn lanes may reduce specific crash types relating to severe consequences such as head-on crashes, while exclusive right-turn lanes may increase some crash types corresponding to less severe injuries such as read-end crashes at sign-controlled intersections. The correlation coefficients from the MVPLN model highlight that the crash counts are highly correlated among all crash severities, which indicates that accounting for their correlations might yield more accurate estimation results when simultaneously estimating crash counts by severity level.

The Joint NB-GOPFS model has two components, where a NB modeling framework is used to estimate the total crash counts, and a GOPFS modeling framework is used to estimate the crash proportions by each severity level. In the NB modeling framework, a positive coefficient indicates a positive correlation between the independent variable and total crash counts, and vice versa. In the GOPFS modeling framework, a positive coefficient represents that the independent variable is associated with increased proportions of severe injury crashes, and vice versa. The coefficient estimates in the NB modeling framework are consistent with the MVPLN model, in which the major and minor road AADT, 4-leg intersections and exclusive right-turn lanes are associated with increased total crash counts, while all-way controlled intersections and presence of driveways are associated with decreased total crash counts. The coefficient estimates of GOPFS modeling framework illustrate that more traffics in major road and 4-leg intersections are highly correlated with increased proportions of severe injury crashes, and exclusive left-turn lanes are significantly associated with decreased proportions of severe injury crashes. The threshold parameters in the Joint NB-GOPFS model indicate the demarcation points between severity categories, which have no substantial interpretation (Yasmin & Eluru, 2018). One important finding in the Joint NB- GOPFS model is that the total crash counts and the threshold between the proportions of B + C and K + A crashes are positively correlated. This finding implies that sites with higher number of total crashes are more likely to incur higher proportions of B + C crashes (as the threshold will move rightward in the generalized ordered Probit fractional split framework, and the thresholds for B + C and K + A are used to define the crash proportions of B + C crashes), and their correlation is found to be constant across different intersections. This verifies the presence of common unobserved factors affecting both total crash counts and the proportions of crashes by severity and accounting for the unobserved factors when simultaneously estimating total crashes and crash severity proportions may provide more accurate estimation results.

4.1.2. Model estimation for vehicle damage component

As mentioned earlier, we also estimated crash prediction models by vehicle damage level to supplement the injury severity, which can be used as an alternative to identify locations that may experience severe injury crashes in the future when the current sample mean of severe injury crashes is very low, which leads to the difficulty of developing crash prediction models by injury severity. As shown in the results, the MVPLN model coefficient estimates regarding the vehicle damage component are highly consistent with the injury severity component. The correlation coeffi-

Table 1
Descriptive Characteristics of Urban & Suburban Intersection and Crash Data.

Variables	Sign-Controlled Intersections (895 Intersections)		Signalized Intersections (1,178 Intersections)	
Crash Data				
K + A Crash	Min. = 0; Max. = 2; Mean = 0.1; Std. Dev = 0.3		Min. = 0; Max. = 5; Mean = 0.2; Std. Dev = 0.5	
B + C Crash	Min. = 0; Max. = 19; Mean = 1.6; Std. Dev = 2.3		Min. = 0; Max. = 54; Mean = 6.2; Std. Dev = 6.3	
PDO Crash	Min. = 0; Max. = 61; Mean = 5.1; Std. Dev = 6.4		Min. = 0; Max. = 146; Mean = 20.3; Std. Dev = 18.4	
Severe Damage Crash	Min. = 0; Max. = 22; Mean = 2.4; Std. Dev = 2.7		Min. = 0; Max. = 60; Mean = 5.9; Std. Dev = 5.1	
Moderate Damage Crash	Min. = 0; Max. = 23; Mean = 1.4; Std. Dev = 2.0		Min. = 0; Max. = 40; Mean = 5.0; Std. Dev = 5.3	
Minor Damage Crash	Min. = 0; Max. = 54; Mean = 3.3; Std. Dev = 4.5		Min. = 0; Max. = 129; Mean = 15.5; Std. Dev = 15.7	
Intersection Data	**Frequency**	**Percentage**	**Frequency**	**Percentage**
3-Leg Intersection	687	76.8%	423	35.9%
4-Leg Intersection	208	23.2%	755	64.1%
Partial-Way Sign-Controlled Intersection	806	90.1%	NA	NA
All-Way Sign-Controlled Intersection	89	9.9%	NA	NA
Median Presence at Intersection Approaches	83	9.3%	285	24.2%
Illumination Presence	599	66.9%	877	74.4%
Driveway Presence	363	40.6%	478	40.6%
Exclusive Left-Turn Lane Presence	51	5.7%	766	65.0%
Exclusive Right-Turn Lane Presence	40	4.5%	586	49.7%
Protected Left-Turn Signal Phasing Presence	NA	NA	875	74.3%
No Right-Turn-On-Red	NA	NA	654	55.5%
Major Road AADT	Min. = 550; Max. = 32,800; Mean = 9,180; Std. Dev = 4,842		Min. = 2,300; Max. = 68,200; Mean = 15,769; Std. Dev = 6,676	
Minor Road AADT	Min. = 20; Max. = 13,900; Mean = 2,534; Std. Dev = 2,212		Min. = 300; Max. = 43,300; Mean = 7,482; Std. Dev = 5,466	
Intersection Skew Angle	Min. = 0; Max. = 89; Mean = 22; Std. Dev = 20		Min. = 0; Max. = 90; Mean = 22; Std. Dev = 23	

cients show that the crash counts by different vehicle damage levels are significantly correlated.

Similarly, the coefficient estimates for the Joint NB-GOPFS model are consistent with those for the injury severity component. The correlation coefficients in the Joint NB-GOPFS demonstrate that the total crash counts and the proportions of crashes by vehicle damage are positively correlated, which implies that sites with a higher number of crashes are more likely to incur higher proportions of severe vehicle damage crashes. The consistent model estimation results between injury severity and vehicle damage components provide support to our initial hypothesis of using vehicle damage as a supplemental indicator of injury severity for estimating crash prediction models by different severity levels.

4.2. Urban & suburban signalized intersections

4.2.1. Model estimation for injury severity component

Table 3 presents the model estimation results for urban & suburban signalized intersections. In terms of the injury severity component, the estimated coefficients for major and minor road AADT and 4-leg intersections are consistent with the urban & suburban sign-controlled intersections, and are associated with increased crash counts for all three levels of crash severity. Presence of driveway is found to be correlated with increased B and C and PDO crashes at signalized intersections. The exclusive right-turn lanes are found to be negatively associated with all severity counts at signalized intersections. One interesting finding is that the protected left-turn signal phasing is correlated with decreased severe injury crashes (K and A crashes), but is correlated with increased B, C and PDO crashes. This might be because the protected left-turn signal phasing can be effective at reducing the head-on crashes, but it might increase the rear-end crashes when the leading vehicle unexpectedly brakes and collided by the following vehicle when the left-turn signal turns to yellow or red. The presence of no right-turn-on-red at signalized intersections is correlated with the increased PDO crashes only, which may be due to the driver's violation of this type of traffic control. The MVPLN model indicates that the crash counts are highly correlated among all crash severity levels at the urban & suburban signalized intersections.

With respect to the Join NB-GOPFS model, the estimation results for total crashes in the NB modeling component are still consistent with the MVPLN model. Three variables are found to be significant for estimating crash proportions by severity level in the GOPFS modeling component. 4-leg intersections are associated with increased proportions of severe injury crashes. If a depressed median is present on any of the intersection approaches, the proportions of severe injury crashes are expected to be increased. The exclusive left-turn lanes are associated with decreased proportions of severe injury crashes. Different from the urban & suburban sign-controlled intersections, the estimated correlation coefficients from the Joint NB-GOPFS model indicate that the total crash counts and crash severity proportions are independent at urban & suburban signalized intersections.

4.2.2. Model estimation for vehicle damage component

The model estimation results for the vehicle damage component yield consistent parameters with the injury severity component, and the crash counts are prone to be correlated among crashes with all vehicle damage levels. For the Joint NB-GOPFS model, higher traffic volumes yield decreased proportions of severe damage crashes, which fits the expectation because vehicle speed tends to be lower when the traffic is heavy. Protected left-turn sig-

Table 2
Model Estimation Results for Sign-Controlled Intersections.

Variables	**Injury Severity Component**				
	MVPLN Model			**Joint NB-GOPFS Model**	
	K + A	**B + C**	**PDO**	**Total Crashes**	**Severity Proportions**
Constant	−17.11 (0.00)	−17.59 (0.00)	−18.24 (0.00)	−9.84 (0.00)	NA
Ln (Major AADT)	0.55 (0.03)	0.68 (0.00)	0.82 (0.00)	0.80 (0.00)	0.10 (0.10)
Ln (Minor AADT)	0.24 (0.05)	0.54 (0.00)	0.60 (0.00)	0.56 (0.00)	–
4-Leg Intersection	0.46 (0.08)	0.77 (0.00)	0.46 (0.00)	0.64 (0.00)	0.11 (0.09)
All-Way Sign-Controlled	−1.21 (0.06)	−0.55 (0.00)	–	−0.31 (0.00)	–
Driveway Presence	−0.45 (0.07)	−0.18 (0.02)	−0.16 (0.00)	−0.16 (0.01)	–
Exclusive Left-Turn Lane Presence	−1.71 (0.05)	−0.41 (0.02)	−0.19 (0.08)	–	−0.23 (0.04)
Exclusive Right-Turn Lane Presence	–	0.59 (0.00)	0.37 (0.00)	0.35 (0.00)	–
Overdispersion	0.65 (0.01)	0.47 (0.00)	0.33 (0.00)	0.24 (0.05)	NA
Threshold 1	NA	NA	NA	NA	1.10 (0.06)
Threshold 2	NA	NA	NA	NA	0.43 (0.00)
Correlation Coefficients	K + A	B + C	PDO	**Correlation Coefficients**	Total Crashes
K + A	1.00	0.79 (0.00)	0.53 (0.00)	Propensity of proportions of severe injury crashes	–
B + C		1.00	0.74 (0.00)	Threshold between B + C and K + A proportions	0.31 (0.09)
PDO			1.00	Threshold between PDO and B + C proportions	–
Variables	**Vehicle Damage Component**				
	MVPLN Model			**Joint NB-GOPFS Model**	
	Severe Damage	**Moderate Damage**	**Minor Damage**	**Total Crashes**	**Damage Proportions**
Constant	−12.64 (0.00)	−16.69 (0.00)	−15.73 (0.00)	−7.10 (0.00)	NA
Ln (Major AADT)	0.35 (0.00)	0.71 (0.00)	0.66 (0.00)	0.65 (0.00)	–
Ln (Minor AADT)	0.35 (0.00)	0.34 (0.00)	0.41 (0.00)	0.39 (0.00)	–
4-Leg Intersection	0.56 (0.00)	0.36 (0.00)	0.16 (0.02)	0.48 (0.00)	0.14 (0.02)
All-Way Sign-Controlled	−0.33 (0.01)	–	–	−0.18 (0.08)	−0.01 (0.09)
Driveway Presence	−0.18 (0.01)	–	–	−0.09 (0.09)	−0.10 (0.07)
Exclusive Right-Turn Lane Presence	0.27 (0.05)	0.75 (0.00)	0.21 (0.08)	0.38 (0.00)	−0.04 (0.04)
Overdispersion	0.33 (0.00)	0.38 (0.00)	0.34 (0.00)	0.04 (0.00)	NA
Threshold 1	NA	NA	NA	NA	−0.96 (0.06)
Threshold 2	NA	NA	NA	NA	−0.69 (0.00)
Correlation Coefficients	Severe Damage	Moderate Damage	Minor Damage	**Correlation Coefficients**	Total Crashes
Severe Damage	1.00	0.73 (0.00)	0.61 (0.00)	Propensity of proportions of severe vehicle damage crashes	0.47 (0.00)
Moderate Damage		1.00	0.78 (0.00)	Threshold between moderate and severe damage proportions	–
Minor Damage			1.00	Threshold between minor and moderate damage proportions	–

Notes: the first value represents the estimated coefficient, followed by the p-value of the coefficient and the following value in parenthesis; "–" represents the variable is not statistically significant at the 10% significance level; "NA" represents the variable is not applicable in the model.

nal phasing is associated with decreased proportions of severe damage crashes. Same as the injury severity component, the total crashes and crash proportions by each vehicle damage level are found to be independent in the Joint NB-GOPFS model.

5. Model comparisons

In order to evaluate the model prediction capability between the MVPLN and Joint NB-GOPFS models, we randomly selected 80% of the datasets (*i.e.*, estimation datasets) to estimate the model coefficients, and used the remaining 20% datasets (*i.e.*, validation datasets) to evaluate the model prediction accuracy, based on the criteria of Mean Absolute Error (MAE) calculated as:

$MAE = \sum_{i=i}^{N} \frac{|Y_{i,pred.} - Y_{i,obs.}|}{N}$ (19)where $Y_{i,pred.}$ represents the predicted crash counts for intersection i corresponding to the specific injury severity or vehicle damage; $Y_{i,obs.}$ represents the observed crash counts for intersection i corresponding to the specific injury severity or vehicle damage; and N represents the sample size.

A smaller *MAE* value indicates a better prediction accuracy. The *MAE* is calculated for both model estimation (*EMAE*) and validation (*VMAE*) datasets. Figs. 2 and 3 present the model prediction comparison results. In general, the MVPLN model performs slightly better in predicting severe crashes, while the Joint NB-GOPFS model performs better in predicting less severe crashes. Specifically, in terms of the crash prediction by injury severity for both sign-controlled and signalized intersections, the MVPLN model slightly outperforms the Joint NB-GOPFS model in predicting K and A crashes, while the Joint NB-GOPFS model performs better in predicting B, C and PDO crashes. The Joint NB-GOPFS model has a smaller prediction error than the MVPLN model based on the average *MAE* value across all severity levels. With regard to the crash prediction by vehicle damage, the MVPLN model performs slightly better than the GOPFS model in predicting severe and minor damage crashes and the average crashes across all damage levels for sign-controlled intersections, but it only outperforms the GOPFS model in predicting severe damage crashes for signalized intersection.

6. Summary and conclusions

This paper presents two advanced frameworks in predicting crash counts by each severity level (*i.e.*, either directly estimating

Table 3
Model Estimation Results for Signalized Intersections.

Variables	**Injury Severity Component**				
	MVPLN Model			**Joint NB-GOPFS Model**	
	K + A	**B + C**	**PDO**	**Total Crashes**	**Severity Proportions**
Constant	−22.12 (0.00)	−17.56 (0.00)	−17.37 (0.00)	−9.58(0.00)	NA
Ln (Major AADT)	0.89 (0.00)	0.81 (0.00)	0.89 (0.00)	0.87 (0.00)	–
Ln (Minor AADT)	0.60 (0.00)	0.40 (0.00)	0.43 (0.00)	0.48 (0.00)	–
4-Leg Intersection	0.24 (0.10)	0.48 (0.00)	0.40 (0.00)	0.43 (0.00)	0.09 (0.00)
Intersection Approach Median Presence	–	–	–	–	0.07 (0.03)
Driveway Presence	–	−0.20 (0.00)	−0.17 (0.00)	−0.17 (0.01)	–
Exclusive Left-Turn Lane Presence	–	–	–	–	−0.09 (0.01)
Exclusive Right-Turn Lane Presence	−0.32 (0.05)	−0.17 (0.00)	−0.10 (0.01)	−0.08 (0.05)	–
Protected Left-Turn Signal Phasing Presence	−0.35 (0.07)	0.09 (0.06)	0.12 (0.00)	–	–
No Right-Turn-On-Red	–	–	0.07 (0.01)	–	–
Overdispersion	0.54 (0.00)	0.34 (0.00)	0.23 (0.00)	0.29 (0.00)	NA
Threshold 1	NA	NA	NA	NA	0.55 (0.14)
Threshold 2	NA	NA	NA	NA	0.56 (0.00)
Correlation Coefficients	K + A	B + C	PDO	**Correlation Coefficients**	Total Crashes
K + A	1.00	0.77 (0.00)	0.65 (0.00)	Propensity of proportions of severe injury crashes	–
B + C		1.00	0.83 (0.00)	Threshold between B + C and K + A proportions	–
PDO			1.00	Threshold between PDO and B + C proportions	–
Variables	**Vehicle Damage Component**				
	MVPLN Model			**Joint NB-GOPFS Model**	
	Severe Damage	**Moderate Damage**	**Minor Damage**	**Total Crashes**	**Damage Proportions**
Constant	−15.32 (0.00)	−19.37 (0.00)	−18.98 (0.00)	−9.53 (0.00)	NA
Ln (Major AADT)	0.67 (0.00)	0.95 (0.00)	0.93 (0.00)	0.86 (0.00)	−0.08 (0.03)
Ln (Minor AADT)	0.33 (0.00)	0.44 (0.00)	0.53 (0.00)	0.47 (0.00)	−0.08 (0.00)
4-Leg Intersection	0.48 (0.00)	0.42 (0.00)	0.42 (0.00)	0.43 (0.00)	–
Driveway Presence	−0.16 (0.00)	−0.15 (0.00)	−0.20 (0.00)	−0.17 (0.00)	–
Exclusive Left-Turn Lane Presence	–	−0.11 (0.05)	–	–	–
Exclusive Right-Turn Lane Presence	−0.22 (0.00)	−0.15 (0.00)	−0.12 (0.00)	−0.09 (0.03)	–
Protected Left-Turn Signal Phasing Presence	−0.08 (0.09)	0.13 (0.03)	0.14 (0.01)	–	−0.10 (0.00)
Overdispersion	0.22 (0.00)	0.32 (0.00)	0.27 (0.00)	0.28 (0.00)	NA
Threshold 1	NA	NA	NA	NA	−1.41 (0.00)
Threshold 2	NA	NA	NA	NA	−0.68 (0.00)
Correlation Coefficients	Severe Damage	Moderate Damage	Minor Damage	**Correlation Coefficients**	Total Crashes
Severe Damage	1.00	0.69 (0.00)	0.67 (0.00)	Propensity of proportions of severe vehicle damage crashes	–
Moderate Damage		1.00	0.83 (0.00)	Threshold between moderate and severe damage proportions	–
Minor Damage			1.00	Threshold between minor and moderate damage proportions	–

crash counts by different severity levels, or first estimating crash counts in total, and then estimating crash severity distributions to combine with the total crashes for crash count prediction by severity). Two advanced methodologies are implemented with regard to each of the frameworks. In terms of the first framework, a MVPLN model is used to simultaneously estimate crash counts by different severity levels, by accounting for their correlations due to the common unobserved factors. In the second framework, a Joint NB-GOPFS model is applied in which a NB modeling component is used to estimate the total crash counts and a GOPFS modeling component is used to estimate the crash proportions by each severity level. The NB and GOPFS modeling components are jointly estimated by accounting for the correlations between total crashes and crash severity proportions due to the common unobserved factors.

Both sign-controlled and signalized intersections at urban & suburban areas are collected from the State of Connecticut and used for model estimation. The estimated coefficients in the MVPLN model show that crash counts are highly correlated among all severity levels for both sign-controlled and signalized intersections, which indicates that accounting for their correlations might yield more accurate estimation results when simultaneously estimating crash counts by severity. The estimation results of the Joint NB-GOPFS model show that the total crashes are significantly correlated with the proportions of B and C crashes at sign-controlled intersections, and their correlations should be accommodated when simultaneously estimating total crash counts and crash proportions by severity. The total crash counts are found to be independent with the crash proportions by severity at signalized intersections.

In addition, we further estimated crash prediction models by vehicle damage level to supplement the injury severity, which can be used as an alternative to identify locations that may experience severe injury crashes in the future when the current sample mean of severe injury crashes (such as K and A crashes) is very low, which leads to the difficulty of developing crash prediction models

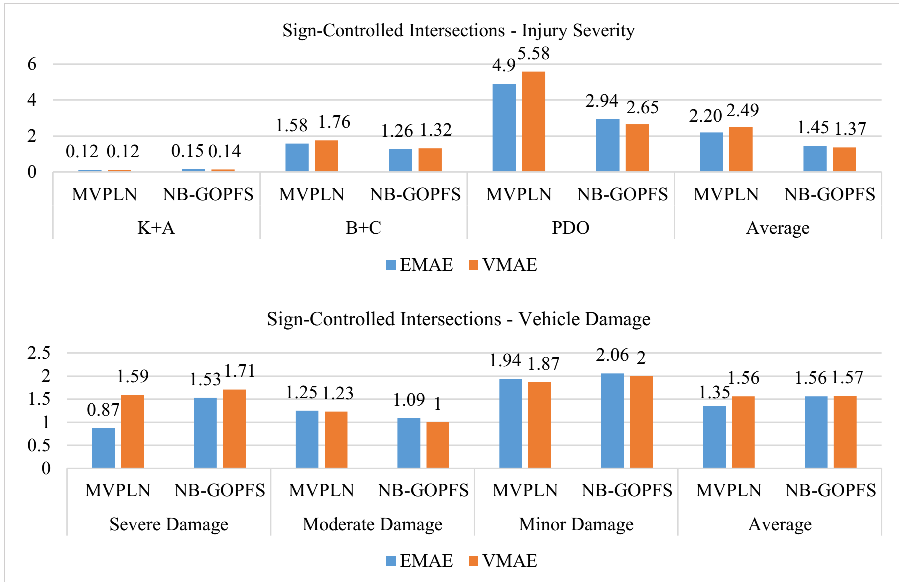


Fig. 2. Model Performance Comparisons for Sign-Controlled Intersections.

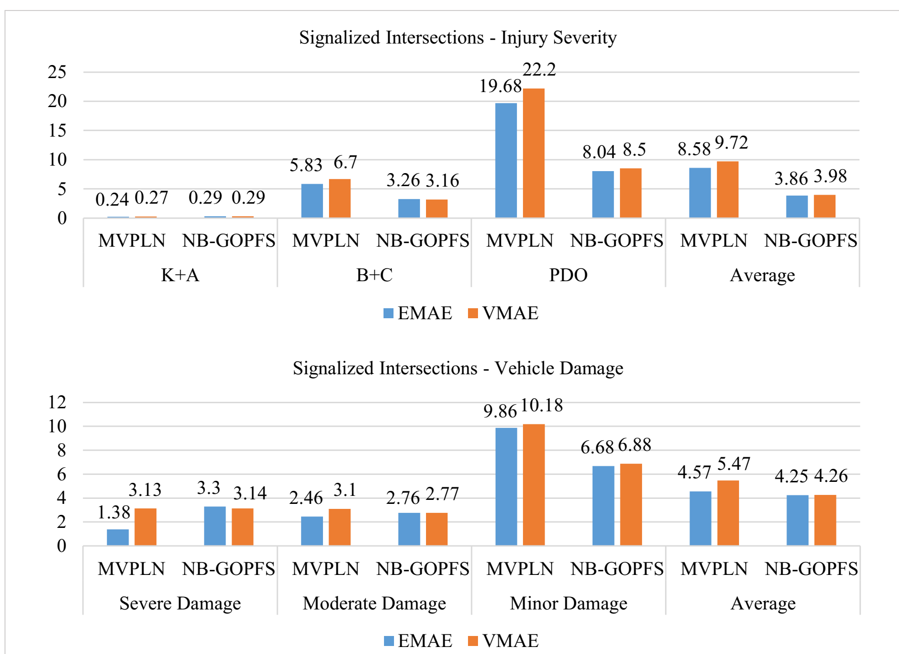


Fig. 3. Model Performance Comparisons for Signalized Intersections.

by injury severity. The model estimation results for injury severity component and vehicle damage component are highly consistent. This finding verifies our initial assumption that when crash data samples have challenges associated with the low observed sampling rates for severe injury crashes, vehicle damage can be appropriate as an alternative to injury severity in crash prediction by severity. An important finding from the model estimation is that two methodologies may yield different variables that are statistically significant in predicting crashes by severity level. For example, the traffic volumes are shown to be significant in all MVPLN models when crash counts by severity are simultaneously modeling, while the traffic volumes seldom affect the prediction of crash proportions by each severity level in the Joint NB-GOPFS model. This may provide additional insight about variable selection in crash prediction models by severity level regarding different approaches. In the end, the prediction performance of the two approaches is compared based on the *MAE* values. The comparisons show that the MVPLN slightly outperforms the Joint NB-GOPFS model in terms of predicting severe crashes, while the Joint NB-GOPFS model significantly improves the prediction accuracy of less severe crashes compared to the MVPLN model. This finding contributes to the practical applications of both crash prediction research and safety improvement effort through shedding light on method selection under different data conditions and research needs, in which the MVPLN is recommended when the analysis target is severe crashes while the NB-GOPFS is preferred for less severe crashes.

7. Practical applications and future work

The findings of this research can offer additional insight into selecting robust methodological modeling frameworks in estimating crash counts by different severity levels, and provide researchers and practitioners with the capabilities of estimating crash prediction models when the low sample mean leads to difficulties in predicting severe crashes. In this study, we used the intersection data to test the proposed modeling frameworks. Future research can focus on extending the modeling frameworks to roadway segments. Future research can also target on extending the MVPLN model to the generalized MVPLN framework, and further extending the Joint NB-GOPFS modeling framework by accounting for the temporal and spatial heterogeneity.

Acknowledgments

The authors would like to thank the Connecticut Department of Transportation (CTDOT) and the Connecticut Crash Data Repository (CTCDR) for providing the intersection and crash data.

Declaration of interest

None.

References

Abdel-Aty, M., & Radwan, A. (2000). Modeling traffic accident occurrence and involvement. *Accident Analysis and Prevention, 32*(5), 633–642.
Aguero-Valverde, J. (2013). Multivariate spatial models of excess crash frequency at area level: Case of Costa Rica. *Accident Analysis and Prevention, 59*, 365–373.
Anarkooli, A., Persaud, B., Hosseinpour, M., & Saleem, T. (2019). Comparison of univariate and two-stage approaches for estimating crash frequency by severity – Case study for horizontal curves on two-lane rural roads. *Accident Analysis and Prevention, 129*, 382–389.
Anastasopoulos, P., Shankar, V., Haddock, J., & Mannering, F. (2012). A Multivariate tobit analysis of highway accident-injury-severity rates. *Accident Analysis and Prevention, 45*, 110–119.
Aptech. Aptech Systems Inc. 2019. https://www.aptech.com/.
Barua, S., El-Basyouny, K., & Islam, M. (2014). A full bayesian multivariate count data model of collision severity with spatial correlation. *Analytic Methods in Accident Research, 3–4*, 28–43.
Barua, S., El-Basyouny, K., & Islam, M. (2016). Multivariate random parameters collision count data models with spatial heterogeneity. *Analytic Methods in Accident Research, 9*, 1–15.
Bhat, C. (2001). Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research Part B: Methodological, 35*(7), 677–693.
Bhowmik, T., Yasmin, S., & Eluru, N. (2019). A multilevel generalized ordered probit fractional split model for analyzing vehicle speed. *Analytic Methods in Accident Research, 21*, 13–31.
Chiou, Y., & Fu, C. (2013). Modeling crash frequency and severity using multinomial generalized poisson model with error components. *Accident Analysis & Prevention, 50*, 73–82.
Chiou, Y., Fu, C., & Chih-Wei, H. (2014). Incorporating spatial dependence in simultaneously modeling crash frequency and severity. *Analytic Methods in Accident Research, 2*, 1–11.
Chiou, Y., & Fu, C. (2015). Modeling crash frequency and severity with spatiotemporal dependence. *Analytic Methods in Accident Research, 5–6*, 43–58.
Connecticut Crash Data Repository. Connecticut Transportation Safety Research Center. 2019. http://www.ctcrash.uconn.edu/.
Dixon, K., Monsere, C., Avelar, R., Barnett, J., Escobar, P., Kothuri, S., & Wang, Y. (2015). *Improved safety performance functions for signalized intersections*. Oregon Department of Transportation.
Dong, C., Clarke, D., Yan, X., Khattak, A., & Huang, B. (2014). Multivariate random-parameter zero-inflated negative binomial regression model: An application to estimate crash frequencies at intersections. *Accident Analysis and Prevention, 70*, 320–329.
Eluru, N., Bhat, C. R., & Hensher, D. A. (2008). A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis and Prevention, 40*(3), 1033–1054.
Geedipally, S., Bonneson, J., Pratt, M., & Lord, D. (2013). Severity distribution functions for freeway segments. *Transportation Research Record, 2398*, 19–27.
Highway Safety Manual (HSM) 1st Edition. American Association of State Highway and Transportation Officials. Washington D.C. 2010.
Liu, C., & Sharma, A. (2018). Using the multivariate spatio-temporal bayesian model to analyze traffic crashes by severity. *Analytic Methods in Accident Research, 17*, 14–31.
Lord, D., & Persaud, B. (2000). Accident prediction models with and without trend: Application of the generalized estimating equations procedure. *Transportation Research Record, 1717*, 102–108.
Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice, 44*(5), 291–305.
Ma, J., & Kockelman, K. (2006). Bayesian multivariate poisson-lognormal regression for models of injury count by severity. *Transportation Research Record 1950*, 24–34.
Ma, J., Kockelman, K., & Damien, P. (2008). A multivariate poisson-lognormal regression model for prediction of crash counts by severity, using bayesian methods. *Accident Analysis and Prevention., 40*(3), 964–975.
Mannering, F., & Bhat, C. (2014). Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research, 1*, 1–22.
Mannering, F., Shankar, V., & Bhat, C. (2016). Unobserved heterogeneity and the statistical analysis of highway accident data. *Analytic Methods in Accident Research, 11*, 1–16.
Model Minimum Uniform Crash Criteria, Fifth Edition. U.S. Department of Transportation, Washington, DC. 2017. https://www.nhtsa.gov/mmucc.
National Highway Traffic Safety Administration (NHTSA). Traffic Safety Facts: 2016 Data. 2018. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812580.
Oh, J., Washington, S., & Choi, K. (2004). Development of accident prediction models for rural highway intersections. *Transportation Research Record 1987*, 18–27.
Park, E. S., & Lord, D. (2007). Multivariate poisson-lognormal models for jointly modeling crash frequency by severity. *Transportation Research Record 2019*, 1–6.
Pei, X., Wong, S., & Sze, N. (2011). A joint-probability approach to crash prediction models. *Accident Analysis and Prevention, 43*(3), 1160–1166.
Qin, X., Sultana, A., Chitturi, M., & Noyce, D. (2013). Developing truck corridor crash severity index. *Transportation Research Record, 2386*, 103–111.
Qin, X., Wang, K., & Cutler, C. (2013). Analysis of crash severity based on vehicle damage and occupant injuries. *Transportation Research Record, 2386*, 95–102.
R-INLA. Bayesian Computing with INLA. http://www.r-inla.org/home.
Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society, 71*(Series B), 319–392.
Russo, F., Busiello, M., & Dell, A. (2016). Safety performance functions for crash severity on undivided rural roads. *Accident Analysis & Prevention, 93*, 75–91.
Savolainen, P., Mannering, F., Lord, D., & Quddus, M. (2011). The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. *Accident Analysis and Prevention, 43*(5), 1666–1676.
Serhiyenko, V., Mamun, S., Ivan, J., & Ravishanker, N. (2016). Fast bayesian inference for modeling multivariate crash counts. *Analytic Methods in Accident Research, 10*, 44–53.
Tarko, A. P., Inerowicz, M., Ramos, J., & Li, W. (2008). Tool with road-level crash prediction for transportation safety planning. *Transportation Research Record, 2083*, 16–25.
Ulfarsson, G., & Shankar, V. (2002). Accident count model based on multiyear cross-sectional roadway data with serial correlation. *Transportation Research Record, 1840*, 193–197.
Wang, C., Quddus, M., & Ison, S. (2011). Predicting accident frequency at their severity levels and its application in site ranking using a two-stage mixed multivariate model. *Accident Analysis and Prevention, 43*(6), 1979–1990.
Wang, K., Ivan, J., Burnicki, A., & Mamun, S. (2017). Predicting local road crashes using socioeconomic and land cover data. *Journal of Transportation Safety & Security., 9*(3), 301–318.
Wang, K., Ivan, J., Ravishanker, N., & Jackson, E. (2017). Multivariate poisson lognormal modeling of crashes by type and severity on rural two lane highways. *Accident Analysis and Prevention., 99*, 6–19.
Wang, K., Zhao, S., & Jackson, E. (2018). Multivariate poisson lognormal modeling of weather-related crashes on freeways. *Transportation Research Record, 2672*(38), 184–198.
Wang, K., Zhao, S., & Jackson, E. (2019). Functional forms of the negative binomial models in safety performance functions for rural two-lane intersections. *Accident Analysis and Prevention., 124*, 193–201.
Wang, K., Yasmin, S., Konduri, K., Eluru, N., & Ivan, J. (2015). Copula based joint model of injury severity and vehicle damage in two-vehicle crashes. *Transportation Research Record, 2514*, 158–166.
Wang, K., Bhowmik, T., Yasmin, S., Zhao, S., Eluru, N., & Eric, J. (2019). Multivariate copula temporal modeling of intersection crash consequence metrics: A joint estimation of injury severity, crash type, vehicle damage and driver error. *Accident Analysis and Prevention.*, 188–197.
Washington, S., Karlaftis, M., & Mannering, F. (2011). *Statistical and econometric methods for transportation data analysis* (2nd ed.). Boca Raton, FL: Chapman and Hall/CRC.
Wang, Y., & Kockelman, K. (2013). A Poisson-lognormal conditional-autoregressive model for multivariate spatial analysis of pedestrian crash counts across neighborhoods. *Accident Analysis and Prevention, 60*, 71–81.
Xie, K., Ozbay, K., & Yang, H. (2019). A multivariate spatial approach to model crash counts by injury severity. *Accident Analysis and Prevention, 122*, 189–198.

Yasmin, S., Eluru, N., Lee, J., & Abdel-Aty, M. (2016). Ordered fractional split approach for aggregate injury severity modeling. *Transportation Research Record, 2583*, 119–126.
Yasmin, S., & Eluru, N. (2018). A joint economic framework for modeling crash counts by severity. *Transportmetrica A: Transport Science, 14*(3), 230–255.
Yasmin, S., & Eluru, N. (2013). Evaluating alternate discrete outcome frameworks for modeling crash injury severity. *Accident Analysis and Prevention, 59*(1), 506–521.
Ye, X., Pendyala, R., Shankar, V., & Konduri, K. (2013). A simultaneous equations model of crash frequency by severity level for freeway sections. *Accident Analysis & Prevention, 57*, 140–149.
Zeng, Q., Huang, H., Pei, X., & Wong, S. (2016). Modeling nonlinear relationship between crash frequency by severity and contributing factors by neural networks. *Analytic Methods in Accident Research, 10*, 12–25.

Dr. Kai Wang is a statistician and transportation safety engineer at the Connecticut Transportation Safety Research Center (CTSRC) at the University of Connecticut. His research interests are on the areas of statistical modeling and application in transportation, transportation economics, risk analysis and simulations in transportation construction decision-making and data mining technologies.

Mr. Tanmoy Bhowmik is a PhD candidate in the Department of Civil, Environmental and Construction Engineering at the University of Central Florida. His research interest is mainly geared towards the formulation and application of discreet choice models for analyzing crash frequency and severity for different crash types at both aggregate and disaggregate levels.

Dr. Shanshan Zhao is a research scientist and project manager working at the Connecticut Transportation Safety Research Center (CTSRC) at the University of Connecticut. Her research interests are on the implementation of highway safety manual, predicting crash occurrence and injury severity using statistical models and machine learning methods, analyzing crash causation and driver behavior, and evaluating the effectiveness of safety countermeasures.

Dr. Naveen Eluru is a Professor in the Department of Civil, Environmental and Construction Engineering. His research is geared towards the formulation and development of discrete choice models that allow for better understanding the behavioral patterns involved in various decision processes. He also focuses on examining the applicability of various econometric modeling approaches for crash frequency models, crash severity models and joint frameworks that accommodate crash frequency and severity.

Dr. Eric Jackson is the director of the Connecticut Transportation Institute and the Connecticut Transportation Safety Research Center (CTSRC) at the University of Connecticut. The CTSRC assisted in the complete overhaul and modernization of crash data and safety analysis in the State. Dr. Jackson's current research has focused on improving the crash data collection process in Connecticut, providing public access to crash data and transportation safety analysis tools, and building the environment for connected and autonomous vehicles in the State.

Journal of Safety Research 76 (2021) 36–43

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Improving knowledge of cyclist crashes based on hospital data including crash descriptions from open text fields

Mette Møller *, Kira Hyldekaer Janstrup, Ninette Pilegaard

Technical University of Denmark, Department of Technology, Management and Economics, Division of Transport, DK-2800 Kgs. Lyngby, Denmark

ARTICLE INFO

Article history:
Received 28 October 2019
Received in revised form 19 June 2020
Accepted 16 November 2020
Available online 10 December 2020

Keywords:
Underreporting
In-depth analysis
Bicycle safety
Crash analysis
Binary logit model

ABSTRACT

Introduction: In this study we explore the added value of bicycle crash descriptions from open text fields in hospital records from the Aarhus municipality in Denmark. We also explore how bicycle crash data from the hospital complements crash data registered by the police in the same area and time period. *Method*: The study includes 5,313 Danish bicycle crashes, of which 4,205 were registered at the hospital and 1,078 by the police. All crashes occurred from 2010 to 2015. We performed an in-depth analysis of the open text fields on hospital records to identify factors associated with each crash using four categories: bicyclist, road, bicycle, and the other party. We employed the chi-squared test to compare the distribution of variables between crashes registered at the hospital and by the police. A binary logit model was used to estimate the probability that a crash factor is identified, and that each crash factor is associated with a single-bicycle crash. *Results*: The open-ended text fields in hospital records provide detailed information about crash factors not available in police records, including riding speed, inattention, clothing, specific road conditions, and bicycle defects. The factors alcohol and curb had the highest odds of being identified in relation to a single-bicycle crash. Crash data registered at the hospital included a larger number of bicycle crashes, particularly single-bicycle crashes and crashes with slight injuries only. *Conclusion*: Crash information registered at the hospital in Aarhus Municipality contributes to a better understanding of bicycle crashes due to detailed information about crash-associated factors as well as information about a larger number of bicycle crashes, particularly single-bicycle crashes. *Practical implication:* Efforts to improve access to detailed information about bicycle crashes are needed to provide a better basis for bicycle crash prevention.

1. Introduction

The individual positive environmental (Xia et al., 2013) and health benefits of bicycling generally outweigh the possible negative effects of increased exposure to air pollution (Woodward & Samet, 2016). These and other positive impacts of bicycling have contributed to a substantial increase in initiatives to promote bicycling over the past decade (Schepers & Heinen, 2013), as bicycling is considered a sustainable means of daily transport and is often mentioned as an important element to ensure a sustainable transport sector. Unfortunately, bicycling is related to a comparably high risk of road traffic injury (Elvik, 2009). In Denmark, the risk of being injured or killed in a bicycle crash is 13 times higher per travelled kilometer than the risk related to car travel (Christiansen & Warnecke, 2018). Research supports the idea of safety in numbers for bicyclists, and thus that the number of bicycle crashes does not increase proportionally with an increasing number of cyclists (Elvik & Bjørnskau, 2017; Jacobsen, 2003). However, changes in bicycling patterns and frequency do lead to changes in the type and severity of crashes (Schepers, Stipdonk, Methorst, & Olivier, 2017). Information on risk factors associated with bicycle crashes remains relevant to develop targeted preventive measures.

Traditionally, knowledge of road traffic crashes relies upon police registered data. However, due to the high level of underreported bicycle crashes in such crash data (e.g. Elvik & Mysen, 1999) information about the occurrence and characteristics of bicycle crashes is incomplete, particularly regarding single-bicycle crashes (e.g., Juhra et al., 2012). In Denmark it is estimated that the level of underreporting regarding bicycle crashes is approximately 90%, as only approximately 10% of bicyclist crashes are registered by the police and thus appear in the official national road traffic crash statistics (Janstrup, Kaplan, Hels, Lauritsen, &

* Corresponding author at: Technical University of Denmark, Department of Technology, Management and Economics, Section of Transport, Bygningstorvet, Building 116B, DK-2800 Kgs. Lyngby, Denmark.

E-mail address: mette@dtu.dk (M. Møller).

https://doi.org/10.1016/j.jsr.2020.11.004
0022-4375/

Prato, 2016). Hospital data do not include all bicycle crashes either, but previous studies have shown that the level of underreporting of bicycle crashes in hospital data is lower than in crash data registered by the police (e.g., Janstrup et al., 2016; Watson, Watson, & Vallmuur, 2015). Further insight into differences and similarities between bicycle crashes registered at a hospital and by the police is needed to fully understand the contribution and limitations of these data.

The Danish national road traffic crash database includes information about a number of factors systematically registered by the police. The registration follows a template that includes aspects such as time of day, weather condition, vehicles involved, crash situation, and crash location. However, information regarding the behavior of the bicyclist, the road conditions and the interaction between the bicyclist and the surroundings is limited and restricted to aspects included in the template. Consequently, a detailed understanding of factors associated with the crash is difficult. Moreover, exploring influence from aspects not included in the template – such as distracted bicycling or an uneven surface – is not possible. To further improve the understanding of bicycle crashes, data sources that include more detailed information about the crash, the surroundings, and the behavior of the bicyclist and allow a more explorative approach is relevant. In line with the conceptual framework for bicycle safety outlined by Schepers, Hagenzieker, Methorst, van Wee, & Wegman (2014), previous studies applying an explorative approach on more detailed data sources (e.g., Møller & Haustein, 2016) have successfully categorized crash associated factors into three categories: (a) the road and its surroundings, (b) the vehicle, and (c) the behavior and condition of the road user. This categorization will also be applied in this study.

Based on the above, this study sets out to explore the added value of bicycle crash descriptions in open text fields in hospital records from Aarhus municipality. The study also explores how bicycle crash data from the hospital complement police registered crash data from the same area and time period, thereby improving the basis for increasing bicycle safety. Exploring crash descriptions in open text fields adds to previous research on bicycle crashes including hospital data (e.g., Langley, Dow, Stephenson, & Kypri, 2003; Lujic, Finch, Boufous, & Hayen, 2008; Juhra et al., 2012; Watson et al., 2015) by facilitating a more detailed understanding of factors associated with a crash.

2. Method

2.1. Road traffic crash registration by the police and at Aarhus municipality

In Denmark, the police only register a bicycle crash if the police are notified about the crash and the crash occurred on a public road. Bicycle crashes occurring in a forest or on a private road are not registered. When notified, the police are only obligated to register the crash if injuries or material damage on a motor vehicle exceeds DKR 50,000 (EUR 6700) or other material damage exceeds DKR 5000 (EUR 670). Due to these criteria, a high number of bicycle crashes are not registered by the Danish police. Regarding injury severity, the police use a scale with four levels: *Fatalities*, the person dies within 30 days of the day of the crash; *severe injuries*, temporary or permanent incapacity; *slight injuries*, the injuries require medical treatment; and *no injuries*, bruises and/or property damage only.

Only a few Danish hospitals register detailed information about road traffic crashes and include open text fields. The hospital in Aarhus Municipality has registered detailed information about road traffic crashes systematically for many years. In addition to open text fields, the hospital follows a template that partly overlaps with the template used by the police, including the four-level injury scale. Road traffic crashes registered at the hospital include crash-involved persons arriving in an ambulance or similar and persons with less severe injuries able to travel to the hospital on their own for a medical check. There are no restrictions regarding crash location for a road traffic crash to be registered at the hospital.

2.2. Road traffic crash data in this study

This study includes all crashes involving bicyclists registered by the police (N = 1,078) and at the hospital (N = 4,205) in the Danish municipality of Aarhus from 2010 to 2015. Both data sources include information on crash characteristics such as day of the week, mode types, crash type (e.g., single-bicycle crash), road user information (gender, age and injury severity) and road conditions (e.g., surface and light conditions). In addition, the hospital data included an open-ended description of the crash containing detailed information about the crash location and cyclist behavior prior to the crash.

2.3. Analysis

To explore the added value of bicycle crash descriptions in open-ended text fields in hospital records, we performed a manual three-step in-depth analysis; first, careful reading of the text; second, identification of crash factors; and third, categorizing the crash factors according to four categories: (a) the behavior and condition of the bicyclist, (b) the road and its surroundings, (c) the bicycle, and (d) the other party. A crash factor is a specific circumstance associated with the crash identified by the bicyclist or hospital staff. Table 1 provides a description of the crash factors regarding the bicyclist, the road and the bicycle. Regarding the category "the other party," the available information indicated that the behavior of the other party was associated with the crash, but the information was too limited to specify a crash factor. We therefore decided to not include those crashes in the detailed analysis of the crash factors. In cases where the description of the crash was missing or the information provided was too limited, a "no crash factor" code was registered.

To explore how bicycle crash data from the hospital complement crash data registered by the police in the same area and time period, we compared the distribution of relevant variables (e.g., injury severity, gender, age group, crash type, light condition, year and season) between the two data sets using the chi-squared test. We then employed a binary logit model to estimate the probability that a crash factor, and a crash factor in a specific category, is identified, as well as the probability that each crash factor was associated with a single-bicycle crash. The binary logit model estimated the probability, based on a function of a vector X_i of observable variables. The observable variables included person and crash characteristics. A significance level of 0.05 was used in all analyses.

3. Results

3.1. Differences between crash data registered at the hospital and by the police

The number of crash-involved bicyclists is significantly higher in the hospital data (N = 4,205, 45%) compared to the police data (N = 1,078, 11%). We also identified significant differences regarding the injury severity of the involved bicyclists. In the hospital data, the percentage of bicyclists with slight injuries (N = 2,410, 57.3%) was significantly higher than in the police data (N = 113,

Table 1
Description of Identified Crash Factors for Bicyclist, Road and Bicycle Categories.

Category	Crash factor	Description
Bicyclist	Alcohol	Cycling under the influence of alcohol
	Inattention	Non-cycling related activities or unspecified inattentiveness
	Bicycling speed	Bicycling speed too high for the conditions or 25+ km/h
	Handling the bicycle	Foot slipping on the pedal, stumbling getting on/off the bicycle, etc.
	Clothing, etc.	Clothing, bags, etc., getting stuck in the wheel
	Crowding	Too little space between the cyclist and other road users
	Violations	Not respecting right of way, red light riding, etc.
	Loss of control	Losing control over the bicycle for no specified reason
	Illness	Acute indisposition
Road	Slippery	Wet leaves, icy, etc.
	Curb	Pedal or wheel hitting the curb
	Design	Bike path too narrow for a bicyclist
	Objects on road	Wire, stone, etc., on the road,
	Road surface	Holes, bumpy surface
	Road works	Hitting or colliding with equipment related to road works
	Weather	Blinded by the weather condition (hard rain, snow, bright sunshine)
	Crossing animal	Crossing cat, etc.
Bicycle	Bicycle chain	Chain breaks
	Various bicycle defects	Saddle comes off, handlebar breaks, etc.
	Brakes	Brakes not working
	Gear	Shift of gears causes turbulence, etc.

10.5%). The police data include a higher percentage of bicyclist fatalities, but the absolute number of bicyclist fatalities is higher in the hospital data (Table 2). The only variable where no significant differences were found was gender, but here the percentage of bicyclists where gender information is missing is higher in the police data (N = 108, 10.0%) than the hospital data (N = 9, 0.2%). In general, the hospital data include a higher percentage of bicyclists in all age groups.

Regarding crash type, the distribution is significantly different in the two data sets. The hospital data include a larger percentage of single-bicycle crashes (N = 2,287, 54.4%) compared to the police data (N = 36, 3.3%), whereas the police data include a larger percentage of intersection crashes. The percentage of crashes occurring in twilight and darkness is higher in the hospital data (31.8% vs. 18.8%), as is the percentage of crashes occurring during the weekend (20.4% vs. 12.7%). Most crashes are registered in the autumn for both data sets, but the distribution across seasons is different.

3.2. Crash factors derived from crash descriptions in open text fields

For 1,274 (30%) of the bicycle crashes registered at the hospital, a crash factor could not be identified due to missing or limited information. Out of the 2,931 crashes with an identified crash factor, 1,025 (33%) regarded the bicyclist, 956 (31%) the road, 96 (3%) the bicycle, and 1,038 (33%) the other party. Table 3 provides an overview of the crash factors identified in the first three categories, as well as the severity degree. In the bicyclist category, alcohol (N = 262, 26%) and inattention (N = 259, 25%) were the most frequently identified crash factors. Clothing, bags, and similar factors were identified in 9% of cases and crowding in 9% of the cases.

Few crash factors were significantly associated with any particular severity degree (Table 3). For crashes where high riding speed (N = 41, 36.6%) or road surface (N = 25, 31.6%) were identified as crash factors, significantly more bicyclists than expected were fatally or severely injured. For crashes where problems with a curb (N = 45, 19.5%) or the road design (N = 41, 33.6%) were identified as crash factors, more cyclists than expected sustained no injury after the crash.

Regarding gender, we found a significant difference in the distribution of crash factors in the bicyclist category ($p > 0.001$), but no significant gender differences in the distribution of crash factors in the two other studied categories (road and bicycle). Regarding the specific crash factors alcohol (N = 184, 70.2%) and bicycling speed (N = 67, 59.8%), these are more common in crashes with a male bicyclist, whereas handling the bicycle (N = 59, 62.8%) and violations (N = 34, 68%) are more frequent crash factors in crashes involving a female bicyclist (Table 4). For the categories inattention, crowding and illness, we found no significant gender differences.

The odds of identifying a crash factor are lower than the odds of not doing so (Table 5). However, compared to crashes with no injury, the probability that a crash factor is identified is higher for injury crashes, particularly for crashes involving severe injury. Regarding crash type, the probability is lowest for single-bicycle crashes and for crashes in "intersections, collision with vehicle from same road." The odds of identifying a crash factor are highest for crashes with "collision with pedestrian or animal" (OR = 3.557), but the 95% confidence interval is high (2.151–5.883) which indicates some uncertainty in the model.

The binary logit model was also employed to see if some crash factors were more likely to be associated with single-bicycle crashes (see Table 6) compared to bicycle crashes involving other parties (vehicle, animal or pedestrian). Some observations were removed from the model due to missing information. The final model is based on information from 2,989 bicyclists of which 1,499 were involved in a single-bicycle crash.

Many of the identified crash factors were significantly associated with the odds of a single-bicycle crash, but the odds ratio values, together with the 95% confidence intervals, show that the uncertainty of these numbers is very high, especially for the variables *slippery*, *objects on the road*, *road surface* and *bicycle chain*. This problem is due to the very low number of non-single-bicycle crashes for which these variables were identified as a crash factor.

Regarding the bicyclist category, we see that the crash factor *clothing* has the highest estimate (5.761), and thus that the probability that the crash is a single-bicycle crash is highest when this crash factor is identified. However, the 95% CI is very wide for this crash factor. The crash factor *alcohol* also has a high estimate (5.421), and the odds of having a single-bicycle crash is more than 200 times higher compared to a crash with another party involved. For this crash factor, the 95% CI is not that big (130.651–391.055). The lowest estimate is found for the crash factor *inattention*

Table 2
Bicycle Crashes Registered by Police and at the Hospital in Aarhus Municipality from 2010 to 2015.

Variable	Category	Police data N	Police data %	Hospital data N	Hospital data %	χ^2-test, p-value
Severity	Fatality	10	0.9	12	0.3	
	Severe injury	209	19.4	832	19.8	
	Slight injury	113	10.5	2410	57.3	
	No injury/N/A	746	69.2	951	22.6	<0.001*
Gender	Male	491	45.5	2115	50.3	
	Female	479	44.5	2081	49.5	
	N/A*	108	10.0	9	0.2	0.9046
Age	0–8 years old	4	0.4	63	1.5	
	9–17 years old	51	4.7	351	8.3	
	18–29 years old	455	42.2	1636	38.9	
	30–45 years old	203	18.9	823	19.6	
	46–65 years old	219	20.3	1083	25.8	
	Older 65 years old	38	3.5	249	5.9	
	N/A*	108	10.0	0	0.0	<0.001
Crash type	Single-bicycle crash	36	3.3	2287	54.4	
	Collision with pedestrian or animal	23	2.1	148	3.5	
	Collision with vehicle on straight road	140	13.0	601	14.3	
	Intersection, collision with vehicle on same road	494	45.8	656	15.6	
	Intersection, collision with vehicle from crossing road	385	35.8	513	12.2	<0.001*
Light condition	Daylight	872	80.9	2827	67.2	
	Twilight	35	3.2	411	9.8	
	Darkness	168	15.6	926	22.0	
	N/A*	3	0.3	41	1.0	<0.001
Surface	Dry	775	71.9	2096	49.8	
	Wet	224	20.8	679	16.1	
	Slippery	23	2.1	378	9.0	
	N/A*	56	5.2	1052	25.1	<0.001
Day	Weekday	941	87.3	3348	79.6	
	Weekend	137	12.7	857	20.4	<0.001*
Season	Winter	185	17.2	807	19.2	
	Spring	247	22.9	1010	24.0	
	Summer	284	26.3	1176	28.0	
	Autumn	362	33.6	1212	28.8	0.008*
Year	2010	130	12.1	616	14.6	
	2011	170	15.8	846	20.1	
	2012	195	18.1	723	17.2	
	2013	208	19.3	757	18.0	
	2014	188	17.4	696	16.6	
	2015	187	17.3	567	13.5	<0.001*

Note: * The N/A category is not included in the statistical test.

(1.425), but the odds of a single-bicycle crash increases by a factor four compared to a crash with another party involved. The small 95% CI (2.751–6.287) for the OR-value for *inattention* shows that the certainty of this number is quite high.

Regarding the road category, the highest probability of a crash being a single-bicycle crash is found when the road is *slippery* (6.314), there is an *object on the road* (6.220), or there are other problems with the *road surface* (6.145). However, the 95% CI for the OR-values of these variables shows a very high uncertainty. Instead the focus should be on the estimates and OR-values for *curb* (5.555, OR = 258.486), *design* (3.624, OR = 37.486), *road works* (5.366, OR = 213.973) and *weather* (3.405, OR = 30.124). The odds of being a single-bicycle crash are highest when the *curb* is identified as a crash factor.

For the bicycle category, the *bicycle chain* has the highest probability (5.960) but the 95% CI is very big and associated with a very high degree of uncertainty.

4. Discussion

The purpose of this study was to explore the added value of bicycle crash descriptions from open text fields in hospital records from Aarhus municipality in Denmark. An additional purpose was to explore how bicycle crash data from the hospital complement crash data registered by the police in the same area and time period. The results suggest that crash data registered at the hospital complement information about bicycle crashes, as these data allow the identification of crash factors not available in the police registered crash data. In addition, hospital data include information about a larger number of bicycle crashes than police crash data, particularly regarding single-bicycle crashes.

The information provided in the open text fields of the medical records allowed detailed insight into factors associated with each crash, such as inattention and specific road conditions. Previous studies have shown that errors are associated with crashes, as well as near crashes, among bicyclists (e.g., Puchades, Pietrantoni, Fraboni, Angelis, & Prati, 2017; Twisk, Commandeur, Vlakveld, Shope, & Kok, 2015). Our results support this and add to the existing literature by identifying specific error types not available for police registered crashes, such as clothing and bags getting stuck in the wheel by mistake, feet slipping on the pedal when trying to get on/off the bicycle, and bicycle defects. The available information did not allow a detailed understanding of the reasons behind the different behaviors associated with each crash, and this therefore remains a relevant topic for future research. In general, our

Table 3
Overview of Crash Factors and Injury Severity of the Involved Bicyclist.

Category	Crash factor	Fatality or severe injury		Slight injury		No injury		Total		χ^2-test,
		N	%	N	%	N*	%	N	%	p-value
Bicyclist	Alcohol	55	21.0	137	52.3	70	26.7	262	100	0.786
	Inattention	60	23.2	148	57.1	51	19.7	259	100	0.120
	Bicycling speed	41	36.6	51	45.5	20	17.9	112	100	<0.001
	Handling the bicycle	24	25.5	56	59.6	14	14.9	94	100	0.061
	Clothing, etc.	15	16.7	55	61.1	20	22.2	90	100	0.454
	Crowding	22	24.7	51	57.3	16	18.0	89	100	0.248
	Violations	6	12.0	27	54.0	17	34.0	50	100	0.207
	Loss of control	10	20.4	25	51.0	14	28.6	49	100	0.853
	Illness[a]	3	15.0	9	45.0	8	40.0	20	100	-
	Total	236	23.0	559	54.5	230	22.4	1025	100	0.011
Road	Slippery road	54	17.5	170	55.2	84	27.3	308	100	0.469
	Curb	40	17.3	146	63.2	45	19.5	231	100	0.034
	Road design	14	11.5	67	54.9	41	33.6	121	100	0.021
	Objects on the road	28	23.9	68	58.1	21	17.9	117	100	0.182
	Road surface	25	31.6	39	49.4	15	19.0	79	100	0.040
	Road works	13	24.5	31	58.5	9	17.0	53	100	0.362
	Weather[a]	4	16.7	14	58.3	6	25.0	25	100	-
	Crossing animal[a]	4	18.2	13	59.1	5	22.7	22	100	-
	Total	182	19.0	548	57.3	226	23.6	956	100	0.002
Bicycle	Bicycle chain	6	14.0	28	65.1	9	20.9	43	100	0.363
	Various bicycle defects	10	32.3	14	45.2	7	22.6	31	100	0.254
	Brakes[a]	6	37.5	8	50.0	2	12.5	16	100	-
	Gear[a]	2	33.3	2	33.3	2	33.3	6	100	-
	Total	24	25.0	52	54.2	20	20.8	96	100	0.1302

[a] Category not included in χ^2-test.

Table 4
Overview of Crash Factors in Bicyclist Category by Gender.

Category	Crash factor	Male		Female		Total		χ^2-test,
		N = 1463	Pct. = 49.9	N = 1461	Pct. = 49.8	N = 2931	Pct. = 100	p-value
Bicyclist	Alcohol	184	70.2	78	29.8	262	100	<0.001
	Inattention	125	48.4	133	51.6	259	100	0.626
	Bicycling speed	67	59.8	45	40.2	112	100	0.042
	Handling the bicycle	35	37.2	59	62.8	94	100	0.015
	Clothing, etc.	45	50.0	45	50.0	90	100	0.995
	Crowding	38	43.2	50	56.8	89	100	0.205
	Violations	16	32.0	34	68.0	50	100	0.014
	Loss of control	24	49.0	25	51.0	49	100	0.884
	Illness	8	40.0	12	60.0	20	100	0.371

Table 5
Factors Influencing Probability a Crash Factor Was Identified.

Variable	Category	Estimate	p-value	OR	95% CI
Intercept		1.122	<0.0001	–	–
Severity	Fatality	−0.934	0.1524	0.393	(0.109–1.412)
	Severe injury	−0.370	0.0008	0.690	(0.556–0.858)
	Slight injury	−0.589	<0.0001	0.555	(0.464–0.663)
	No injury/NA	–	–	–	–
Type	Single-bicycle crash	–	–	–	–
	Collision with pedestrian or animal	1.269	<0.0001	3.557	(2.151–5.883)
	Collision with vehicle on straight road	0.334	0.0013	1.398	(1.140–1.714)
	Intersection, collision with vehicle on same road	0.164	0.0944	1.178	(0.972–1.427)
	Intersection, collision with vehicle on crossing road	0.343	0.0021	1.409	(1.133–1.753)
Log-likelihood		−2491			
AIC		4998			
Pseudo R-square		0.110			

Note: 59 observations were deleted due to missing information.

results support the need for measures to reduce unsafe bicyclist behavior and ensure a bicycle-friendly environment that reduces errors and related consequences in case of a crash, as suggested by the safe systems approach (Wegman, Zhang, & Dijkstra, 2012).

Regarding bicyclist behavior, our results confirm the results of previous studies (e.g. Billot-Grasset, Amoros, & Hours, 2016), which identified bicyclist behavior as a key factor in bicycle crashes. Alcohol impairment and inattention were the most fre-

Table 6
Odds of a Single-Bicycle Crash Compared to Bicycle Crashes involving Several Parties for Each Crash Factor.

Variable	Category	Estimate	p-value	OR	95% CI
Intercept		−2.939	<0.0001	–	–
Bicyclist	Alcohol	5.421	<0.0001	226.035	(130.651–391.055)
	Inattention	1.425	<0.0001	4.159	(2.751–6.287)
	Bicycling speed	3.364	<0.0001	28.890	(17.767–46.978)
	Handling the cycle	4.575	<0.0001	97.063	(52.601–179.107)
	Clothing, etc.	5.761	<0.0001	317.639	(124.094–813.046)
	Loss of control	4.165	<0.0001	64.374	(30.338–136.595)
Road	Slippery	6.314	<0.0001	552.183	(278.889->999.999)
	Curb	5.555	<0.0001	258.486	(139.067–480.45)
	Design	3.624	<0.0001	37.486	(23.157–60.683)
	Objects on the road	6.220	<0.0001	502.536	(179.026->999.999)
	Road surface	6.145	<0.0001	466.597	(142.819->999.999)
	Road works	5.366	<0.0001	213.973	(74.489–614.646)
	Weather	3.405	<0.0001	30.124	(12.617–71.921)
Bicycle	Bicycle chain	5.960	<0.0001	387.575	(91.504->999.999)
	Other defects	5.604	<0.0001	271.450	(81.805–900.74)
Log-likelihood		−770			
AIC		1572			
Pseudo R-square		0.889			

Note: 1307 observations were deleted due to missing information.

quently identified factors. Like Orsi, Ferraro, Montomoli, Otte, and Morandi (2014), we found that female bicyclists were less likely to be influenced by alcohol at the time of the crash, whereas inattention was more prevalent among female bicyclists. Alcohol is known to decrease the ability to behave safely in traffic due to factors such as decreased reaction times, increased error rates, tunnel vision, and slower visual information processing (e.g., Friedman, Robinson, & Yelland, 2011). Evidence has also been established for the harmful effects of being inattentive to bicycling tasks (e.g., De Waard, Edlinger, & Brookhuis, 2011; Terzano, 2013). Our results indicate that efforts to prevent drunk bicycling, as previously suggested by Andersson and Bunketorp (2002), are still highly relevant, particularly for male bicyclists. The data do not allow conclusions regarding the reasons behind drunk cycling; however, underestimation of personal risk has previously been identified as a contributing factor (e.g., Hagemeister & Kronmaier, 2017).

Regarding the environment, a few previous studies have found that roads with poor surface conditions may lead to traffic disruption and increased crash risk (e.g., Corazza, Di Mascio, & Moretti, 2016; Janstrup, Møller, & Pilegaard, 2019; Pulugurtha, Ogunro, Pando, Patel, & Bonsu, 2013). Our results add to this by showing that aspects such as road work sites and slippery and/or bumpy surfaces are also associated with crash involvement among bicyclists. The data do not allow us to draw conclusions regarding specific causal effects; however, in line with the safe systems approach (e.g., Wegman et al., 2012), the data do highlight the importance of designing road environments that support safe behaviors and allow bicyclists to compensate safely in case of human error. Bicycle defects were only identified as a crash factor in a small number of crashes, but nevertheless indicate that better bicycle maintenance has the potential to improve bicyclists' safety.

Our results indicate that females are involved in bicycle crashes related to some type of violation (e.g., no lights, illegal manoeuver) at the time of the crash more frequently than males. This is somewhat surprising, as male road users are generally found to engage in more traffic violations (e.g., Varet, Granié, & Apostolidis, 2018), and a previous study found male bicyclists to be less compliant with road traffic rules (Johnson et al., 2011). It is possible, however, that the gender difference regarding violations is partly related to the data registration procedure. In general, persons identifying with the feminine gender role expectations are more prone to guilt (Benetti-McQuoid & Bursik, 2005; Efthim et al., 2001). This may cause female cyclists to be more willing to admit engaging in traffic violations compared to male cyclists. However, additional studies looking specifically into gender-based differences in bicycling skills, trip purpose, gender role expectations, and other aspects that may explain or contribute to the identified gender differences are relevant, as limited information about gender differences and bicyclist safety exist (Stipancic, Zangenehpour, Miranda-Moreno, Saunier, & Granié, 2016). Our results show that information about gender is more complete in hospital data, which is thus a relevant data source for studies on gender differences in bicycling crashes.

Significantly more crash-involved bicyclists were registered in the hospital data than the police data, thereby confirming that bicycle crash data from the police suffer from a high level of underreporting (e.g., Alsop & Langley, 2001; Langley et al., 2003), and that both data types are relevant for improving bicyclist safety (e.g., Short & Caulfield, 2014). This is particularly true for single-bicycle crashes and crashes involving only slight injuries. Half of the bicyclists registered at the hospital were injured in single-bicycle crashes. This is in line with results from previous studies (e.g., Beck et al., 2016), although the share of single-bicycle crashes in our results was lower than elsewhere (e.g., Schepers, Agerholm et al., 2014). Importantly, our results also showed that, for some crash types, the number of crashes registered by the police was higher than the crash data registered at the hospital. Thus, both data sources are needed to get a complete picture of the prevalence of bicyclist crashes, as neither source is complete.

In our study, the possibility of identifying a crash factor strongly depended on the information provided by the hospital. The results showed that the level of detail provided decreased with increasing crash severity. The data did not allow conclusions regarding the reasons behind this, but it is possible that the severely injured bicyclists' need for immediate care allowed hospital staff less time to focus on gaining a description of the crash situation. Nevertheless, if trying to gain a detailed understanding of the factors associated with bicycle crashes, our results indicate crash data registered at a hospital may be more suitable for less severe crashes than for police registered crashes. However, the probability of identifying a crash factor was lower for single-bicycle crashes compared to all other crash types, which indicates that although more crash details are available in hospital data than police data, efforts are needed to ensure that more information about these and other types of bicycle crashes is registered. In addition, initiatives to ensure that crash information is made available for road safety researchers and practitioners are relevant to improve road safety for cyclists.

5. Limitations

National hospital data regarding cyclist crashes are not available in Denmark. The data included in this study were therefore limited to Aarhus municipality, one of the few Danish municipalities that registered bicycle crash information at the hospital for many years. To validate the results obtained in the present study on a national level, studies including data from a larger share of Danish municipalities would be relevant, but not possible at the moment. Assessing the degree to which the results are transferrable to other countries is difficult; first, because access to the information included is generally highly restricted, and second, because data collection procedures most likely differ across countries. However, similar in-depth analysis would be relevant in other countries as well.

As in other studies (Imprialou & Quddus, 2017), our results are based on information registered by the hospital staff, which has not been verified for accuracy. It has been shown that there is some inconsistency between road traffic crash information registered by the police and by hospital staff (e.g., Lopez, Rosman, Jelinek, Wilkes, & Sprivulis, 2000), partly due to the different purpose for which the data are collected (e.g., civil claims and trauma treatment), which may influence the reporting (Imprialou & Quddus, 2017). Furthermore, like other studies including qualitative data (see Aust, Fagerlind, & Sagberg, 2012; Møller & Haustein, 2016), some variation was found in the type of information and the level of detail provided. These variations may stem from several known and unknown sources (e.g., time pressure or misinterpretations on the part of the hospital staff and social desirability; Lajunen, Carry, Summala, & Harley, 1997) on the part of the injured bicyclist, influencing him or her to leave out certain types of information and prioritize the inclusion of others. However, it was not possible to control for this influence in the present study.

6. Conclusion

This study showed that crash data registered at the hospital contribute significantly to our understanding of bicyclist crashes both in terms of the number of crashes and associated factors. The results showed that preventive measures to reduce inappropriate bicyclist behavior and poor road conditions are of key importance to improve bicyclist safety. Measures to ensure well-maintained bicycles are less crucial but are also relevant for crash and injury prevention. This study confirmed that more bicyclist crashes are registered at the hospital than by the police and that hospital data are important to improve information about bicycle crashes. As hospital data and police data are reported differently, both are necessary to increase the understanding of crash factors and improve crash prevention efforts.

Acknowledgements

Financial support from the Danish Asphalt Pavement Association and the Danish Safe Roads Association is gratefully acknowledged. The authors would like to thank the Municipality of Aarhus for support regarding the data, and the reviewers for their comments, which significantly improved the paper.

References

Alsop, J., & Langley, J. (2001). Under-reporting of motor vehicle traffic crash victims in New Zealand. *Accident Analysis & Prevention, 33*, 353–359.
Andersson, A.-L., & Bunketorp, O. (2002). Cycling and alcohol. *Injury, 33*, 467–471.
Aust, M. L., Fagerlind, H., & Sagberg, F. (2012). Fatal intersection crashes in Norway: Patterns in contributing factors and data collection challenges. *Accident Analysis & Prevention, 45*, 782–791.
Beck, B., Stevenson, M., Newstead, S., Cameron, P., Judson, R., Edwards, E. R., Johnson, M., ... Gabbe, B. (2016). Bicycling crash characteristics: An in-depth crash investigation study. *Accident Analysis & Prevention, 96*, 219–227.
Benetti-McQuoid, J., & Bursik, K. (2005). Individual differences in experiences of and responses to guilt and shame: Examining the lenses of gender and gender role. *Sex Roles, 53*, 133–142.
Billot-Grasset, A., Amoros, E., & Hours, M. (2016). How cyclist behavior affects bicycle accident configurations? *Transportation Research Part F, 41*, 261–276.
Christiansen, H., & Warnecke, M.L. (2018). Risiko i trafikken 2007-2016. DTU-report 2018. http://orbit.dtu.dk/files/146185353/Risiko_i_trafikken_2007_16.pdf.
Corazza, M. V., Di Mascio, P., & Moretti, L. (2016). Managing sidewalk pavement maintenance: A case study to increase pedestrian safety. *Journal of Traffic and Transportation Engineering (English Edition), 3*, 203–214.
De Waard, D., Edlinger, K., & Brookhuis, K. (2011). Effects of listening to music, and of using a handheld and handsfree telephone on cycling behavior. *Transportation Research Part F, 14*, 626–637.
Efthim, P.W., Kenny, M.E., & Mahalik, J.R. (2001). Gender role stress in relation to shame, guilt, and externalization. Journal of Counselling & Development, 79, 46–54.
Elvik, R. (2009). The non-linearity of risk and the promotion of environmentally sustainable transport. *Accident Analysis & Prevention, 41*, 849–855.
Elvik, R., & Bjørnskau, T. (2017). Safety-in-numbers: A systematic review and meta-analysis of evidence. *Safety Science, 92*, 274–282.
Elvik, R., & Mysen, A. (1999). Incomplete accident reporting: Meta-analysis of studies made in 13 Countries. *Transportation Research Record, 1665*, 133–140.
Friedman, T. W., Robinson, S. R., & Yelland, G. W. (2011). Impaired perceptual judgment at low blood alcohol concentrations. *Alcohol, 45*, 711–718.
Hagemeister, C., & Kronmaier, M. (2017). Alcohol consumption and cycling in contrast to driving. *Accident Analysis & Prevention, 105*, 102–108.
Imprialou, M., & Quddus, M. (2017). Crash data quality for road safety research: Current state and future directions. *Accident Analysis & Prevention, 130*, 84–90.
Jacobsen, P. L. (2003). Safety in numbers: More walkers and bicyclists, safer walking and bicycling. *Injury Prevention, 9*, 205–209.
Johnson, M., Newstead, S., Charlton, J., & Oxley, J. (2011). Riding through red lights: The rate, characteristics and risk factors of non-compliant urban commuter cyclists. *Accident Analysis and Prevention, 43*, 323–328.
Janstrup, K. H., Kaplan, S., Hels, T., Lauritsen, J., & Prato, C. G. (2016). Understanding traffic crash under-reporting: Linking police and medical records to individual and crash characteristics. *Traffic Injury Prevention, 17*, 580–584.
Janstrup, K. H., Møller, M., & Pilegaard, N. (2019). A clustering approach to integrate traffic safety in road maintenance prioritization. *Traffic Injury Prevention, 20*, 442–448.
Juhra, C., Wieskötter, B., Chu, K., Trost, L., Weiss, U., Messerschmidt, M., ... Raschke, M. (2012). Bicycle accidents – Do we only see the tip of the iceberg? *Injury, 43*, 2026–2034.
Lajunen, T., Carry, A., Summala, H., & Harley, L. (1997). Impression management and self-deception in traffic behavior inventories. *Journal of Personalized Medicine, 22*, 341–353.
Langley, J. D., Dow, N., Stephenson, S., & Kypri, K. (2003). Missing cyclists. *Injury Prevention, 9*, 376–379.
Lopez, D. G., Rosman, D. L., Jelinek, G. A., Wilkes, G. J., & Sprivulis, P. C. (2000). Complementing police road-crash records with trauma registry data — an initial evaluation. *Accident Analysis & Prevention, 32*, 771–777.
Lujic, S., Finch, C., Boufous, S., & Hayen, A. (2008). How comparable are road traffic crash cases in hospital admissions data and police records? An examination of data linkage rates. Aust N Z J Public Health, 32, 28–33.
Møller, M., & Haustein, S. (2016). Factors contributing to young moped rider accidents in Denmark. *Accident Analysis & Prevention, 87*, 1–7.
Orsi, C., Ferraro, O. E., Montomoli, C., Otte, D., & Morandi, A. (2014). Alcohol consumption, helmet use and head trauma in cycling collisions in Germany. *Accident Analysis & Prevention, 65*, 97–104.
Puchades, V. M., Pietrantoni, L., Fraboni, F., Angelis, M. D., & Prati, G. (2017). Unsafe cycling behaviors and near crashes among Italian cyclists. International journal of injury control and safety promotion. *International Journal of Injury Control and Safety Promotion, 25*, 70–77. https://doi.org/10.1080/17457300.2017.1341931.
Pulugurtha, S. S., Ogunro, V., Pando, M. A., Patel, K. J., & Bonsu, A. (2013). Preliminary results towards developing thresholds for pavement condition maintenance: Safety perspective. *Procedia - Social and Behavioral Sciences, 104*, 302–311.
Schepers, P., Agerholm, N., Amoros, E., Benington, R., Bjørnskau, T., Dhondt, S., ... Niska, A. (2014). An international review of the frequency of single-bicycle crashes (SBCs) and their relation to bicycle modal share. *Inj. Prev., 21*, e138–e143.
Schepers, P., Hagenzieker, M., Methorst, R., van Wee, B., & Wegman, F. (2014). A conceptual framework for road safety and mobility applied to cycling safety. *Accident Analysis and Prevention, 62*, 331–340.
Schepers, P., & Heinen, E. (2013). How does a modal shift from car trips to cycling affect road safety? *Accident Analysis and Prevention, 50*, 1118–1127.
Schepers, P., Stipdonk, H., Methorst, R., & Olivier, J. (2017). Bicycle fatalities: Trends in crashes with and without motor vehicles in The Netherlands. *Transportation Research Part F: Traffic Psychology and Behaviour, 46*, 491–499.
Short, J., & Caulfield, B. (2014). The safety challenge of increased cycling. *Transport Policy, 33*, 154–165.

Stipancic, J., Zangenehpour, S., Miranda-Moreno, L., Saunier, N., & Granié, M.-A. (2016). Investigating the gender differences on bicycle-vehicle conflicts at urban intersections using an ordered logit methodology. *Accident Analysis & Prevention, 97*, 19–27.
Terzano, K. (2013). Bicycling safety and distracted behavior in The Hague, the Netherlands. *Accident Analysis & Prevention, 57*, 87–90.
Twisk, D. A. M., Commandeur, J. J. F., Vlakveld, W. P., Shope, J. T., & Kok, G. (2015). Relationships amongst psychological determinants, risk behaviour, and road crashes of young adolescent pedestrians and cyclists: Implications for road safety education programmes. *Transportation Research Part F: Traffic Psychology and Behaviour, 30*, 45–56.
Varet, F., Granié, M.-A., & Apostolidis, T. (2018). The role of individualism, gender and situational factors on probabilities of committing offences in a French drivers sample. *Transportation Research Part F: Traffic Psychology and Behaviour, 56*, 293–305.
Watson, A., Watson, B., & Vallmuur, K. (2015). Estimating under-reporting of road crash injuries to police using multiple linked data collections. *Accident Analysis & Prevention, 83*, 18–25.
Wegman, F., Zhang, F., & Dijkstra, A. (2012). How to make more cycling good for road safety. *Accident Analysis and Prevention, 44*, 19–29.
Woodward, A., & Samet, J. (2016). Active transport: Exercise trumps air pollution, almost always. *Preventive Medicine, 87*, 237–238.
Xia, T., Nitschke, M., Zhang, Y., Shah, P., Crabb, S., & Hansen, A. (2013). Traffic-related air pollution and health co-benefits of alternative transport in Adelaide, South Australia. *Environment International, 74*, 281–290.

Mette Møller is a Senior Researcher at the Technical University of Denmark. Her research field is traffic safety and transport behaviour. She holds a PhD in social psychology from University of Copenhagen.

Kira h. Janstrup is an Assistant Professor at the Technical University of Denmark. Her research field is road safety and road safety data. She holds a PhD in traffic safety from Technical University of Denmark.

Ninette Pilegaard is a Senior Researcher in transport economics and Deputy Head of the Transport Division at the Technical University of Denmark. She holds a Ph.D. in economics from University of Copenhagen. Her research field is transport economics and welfare appraisal of transport projects and policies.

Journal of Safety Research 76 (2021) 30–35

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Injury-severity analysis of lane change crashes involving commercial motor vehicles on interstate highways

Emmanuel Kofi Adanu [a,*], Abhay Lidbe [a], Elsa Tedla [a], Steven Jones [b]

[a] Alabama Transportation Institute, The University of Alabama Tuscaloosa, AL, United States
[b] Department of Civil, Construction and Environmental Engineering, The University of Alabama Tuscaloosa, AL, United States

ARTICLE INFO

Article history:
Received 13 March 2019
Received in revised form 17 June 2020
Accepted 9 November 2020
Available online 1 December 2020

Keywords:
lane change
Crash severity
Commercial motor vehicles
interstate crashes

ABSTRACT

Introduction: One of the challenging tasks for drivers is the ability to change lanes around large commercial motor vehicles. Lane changing is often characterized by speed, and crashes that occur due to unsafe lane changes can have serious consequences. Considering the economic importance of commercial trucks, ensuring the safety, security, and resilience of freight transportation is of paramount concern to the United States Department of Transportation and other stakeholders. *Method:* In this study, a mixed (random parameters) logit model was developed to better understand the relationship between crash factors and associated injury severities of commercial vehicle crashes involving lane change on interstate highways. The study was based on 2009–2016 crash data from Alabama. *Results:* Preliminary data analysis showed that about 4% of the observed crashes were major injury crashes and drivers of commercial motor vehicles were at-fault in more than half of the crashes. Acknowledging potential crash data limitations, the model estimation results reveal that there is increased probability of major injury when lane change crashes occurred on dark unlit portions of interstates and involve older drivers, at-fault commercial vehicle drivers, and female drivers. The results further show that lane change crashes that occurred on interstates with higher number of travel lanes were less likely to have major injury outcomes. *Practical Applications:* These findings can help policy makers and state transportation agencies increase awareness on the hazards of changing lanes in the immediate vicinity and driving in the blind spots of large commercial motor vehicles. Additionally, law enforcement efforts may be intensified during times and locations of increased unsafe lane changing activities. These findings may also be useful in commercial vehicle driver training and driver licensing programs.

1. Introduction

Safe and efficient freight movement is critical to sustain the economic development of a nation. Based on the 2017 Commodity Flow Survey, commercial motor vehicles (CMVs) account for the majority of freight shipments in the United States – 71.5% tonnage and 73% value (USDOT, 2019). Considering the economic importance of commercial vehicles, ensuring the safety, security, and resilience of freight transportation is of paramount concern to the United States and economies throughout the world. National crash statistics show that a total of 4,440 CMVs were involved in fatal crashes in 2016 and 32% of these crashes occurred on interstates and freeways (NHTSA, 2017; IIHS, 2018). The Federal Highway Administration (FHWA, 2017) estimates the economic cost of fatal CMV crashes to be more than $20 billion each year. As such, there have been a myriad of previous studies into the occurrence (frequency) and outcome (severity) of crashes involving CMVs (e.g., Garber et al., 1992; Braver et al., 1996; Duncan et al., 1998; Chang & Mannering, 1999; Zhu & Srinivasan, 2011; Islam & Hernandez, 2013; Venkataraman et al., 2013; Wang & Shi, 2013; Islam et al., 2014; Pahukula et al., 2015). Zhu and Srinivasan (2011) reported CMV crashes are generally attributable to a range of factors including their proportion of traffic stream and operational differences between them and other smaller vehicles such as visibility, braking, and overall maneuverability.

The physical and operational characteristics of CMVs affect driving activities such as lane changing. Lane changing poses a major potential hazard for many drivers (both CMVs and non CMVs) as the maneuver may require significant cognitive processing and physically interrupting the flow of traffic in one or more lanes (Antin et al. 2017). At higher speeds, as it is on interstates and freeways, the potential hazard is increased. Although not a

* Corresponding author.
E-mail addresses: ekadanu@crimson.ua.edu (E.K. Adanu), adlidbe@ua.edu (A. Lidbe), egtedla@eng.ua.edu (E. Tedla), sjones@eng.ua.edu (S. Jones).

https://doi.org/10.1016/j.jsr.2020.11.001
0022-4375/

primary reason for a lane change maneuver, CMVs can generate wind gusts that can push smaller vehicles into other lanes, forcing an unintentional lane change (Federal Motor Carrier Safety Administration, 2018). Indeed, it is estimated that more than 250,000 crashes occur annually in the United States because of lane changing errors (Sen et al., 2003). The ability to successfully execute a lane change is therefore influenced by many factors, including driver and vehicle characteristics, roadway features and conditions, and traffic conditions. Due to its peculiarity and potential for serious injuries, this paper presents an analysis of factors affecting severity outcomes of lane change crashes involving CMVs on interstates in Alabama. A mixed logit modeling technique was adopted to address the inherent problem of unobserved heterogeneity in crash data. Mixed logit has also been shown to be statistically superior to the traditional fixed parameters logit models (e.g., Anastasopoulos & Mannering, 2011; Chen & Chen, 2011; Morgan & Mannering, 2011; Islam et al., 2014).

2. Methodology

Unobserved heterogeneity is an important concern in traffic safety research. Failure to account for the effect of unobserved variables can lead to biased estimates and incorrect inferences if inappropriate methods are used (Mannering et al., 2016). Discrete-outcome models, ordered (probit and logit models), and unordered models (such as the multinomial logit model) have been used to analyze crash severity due to the classification of the severities into discrete outcomes (see Savolainen et al., 2011; Mannering & Bhat, 2014 for injury-severity methodology reviews). However, many of the discrete ordered and unordered models are unable to account for unobserved heterogeneity across injury-severity observations (see Savolainen et al., 2011). To capture the effects of unobserved heterogeneity due to randomness associated with some of the factors necessary to understand crash injury severity, this study used mixed logit modeling techniques. Details of the mixed logit model formulation are summarized in Table 1.

Mixed logit model is a generalization of the multinomial logit model. Unlike the multinomial logit model, the mixed logit model allows the parameter vector β_i to vary across each observation in a manner that the injury severity specific constant and each element of the parameter vector β_i can be fixed or randomly distributed with fixed means, where the mixed logit weights for these random parameters are determined by a known density function $f(\beta_i|\varphi)$. Statistically significant variance in φ is an indication that the modeled injury severity varies with respect to Xacross individual observations as defined by $f(\beta_i|\varphi)$ (Train, 2003). Many studies have used normal distribution in model estimation (Milton et al., 2008). Though recent studies have shown that mixed logit models with heterogeneity in means and mixed logit models with heterogeneity in means and variances perform better than the conventional mixed logit models (e.g. Greene et al., 2006; Venkataraman et al., 2014; Kim et al., 2013; Seraneeprakarn et al., 2017; Behnood & Mannering, 2017), the conventional mixed logit model was used for this study because it has demonstrated a good fit of the data.

Table 1
Equations used in mixed logit model formulation.

Equation	Description
$S_{in} = \beta_i X_{in} + \varepsilon_{in}$	S_{in} = severity function for category i in crash n β_i = estimable severity parameters for category i, X_{in} = explanatory variables of severity category i in crash n, ε_{in} = error term – generalized extreme value distributed (McFadden, 1981)
$P_n(i) = \frac{\exp(\beta_i X_{in})}{\Sigma exp(\beta_i X_{in})}$	$P_n(i)$ = probability of ith outcome occurring in the nth observation (Washington et al., 2011)
$P_n(i\|\varphi) = \int \frac{exp(\beta_i X_{in})}{\Sigma exp(\beta_i X_{in})} f(\beta_i\|\varphi)\beta_i$	$P_n(i\|\varphi)$ = probability of injury severity i conditional on $f(\beta_i\|\varphi)$ φ = vector of parameters with known density function (McFadden and Train, 2000; Train, 2003)

Maximum likelihood estimation of mixed logit models is computationally complex. This is because of the difficulty associated with the required numerical integration of the logit function over the distribution of random, unobserved parameters. Simulation based methods are therefore generally used. The Halton sequence (or Halton draws), based on a methodology developed by Halton (1960) to generate a systematic non-random sequence numbers, is the widely used approach to achieve this. Bhat (2003) and Train (1999) have shown Halton draws to be more efficient and also achieve accurate probability approximations with less draws, than purely random draws (Train, 1999; Bhat, 2003). As recommended by Bhat (2003), Halton draws (200) was used to estimate possible mixing distributions for this study.

Elasticities are commonly computed from the partial derivative for each observation to investigate the effect of individual parameters on the injury-severity outcome probabilities (Washington et al., 2011) as:

$$E_{x_{ki}}^{P(i|\varphi)} = \frac{\partial P(i|\varphi)}{\partial x_{ki}} \times \frac{x_{ki}}{P(i|\varphi)}$$

where, $P(i|\varphi)$ is the probability of injury-severity outcome i and x_{ki} is the value of variable k for outcome i. A pseudo-elasticity can also be computed for indicator variables to show the percent effect of the variable taking a value of zero to one on the injury-severity outcome probability (e.g., Islam and Jones, 2014; Islam et al., 2014; Shaheed et al., 2013). For example, a 0.12 average direct pseudo-elasticity of a variable on a probability indicates that the probability increases by 12% on average when the variable is changed from 0 to 1.

3. Data and empirical setting

The study was based on 2009–2016 crash data obtained from the Critical Analysis Reporting Environment (CARE) system developed by the Center for Advanced Public Safety at the University of Alabama. This is the primary database where crash records input directly by all traffic safety law enforcement officers in the State of Alabama are maintained. Each year the data goes through a rigorous QA/QC process consistent with typical traffic safety databases maintained by state agencies throughout the United States. The CARE database serves as the primary source of historical crash data for research and policy decision-making in the State of Alabama. The database was queried to select interstate crashes in which the primary contributing factor was unsafe lane change and involved at least one CMV. CARE defines commercial vehicle as a motor vehicle designed or used to transport passengers or property and meeting at least one of the following criteria:

a) Vehicle has a Gross Vehicle Weight Rating (GVWR) of 26,001 or more pounds, or
b) Vehicle is designed to transport 16 or more passengers, including the driver, or
c) Vehicle is transporting hazardous materials and is required to be placarded.

The need to accurately establish responsibility for crashes is well documented (e.g., Carr, 1969; Roberston & Drummer, 1994) and new methods for doing so continue to be developed and tested (e.g., Salmi et al., 2012; Garcia et al., 2019). Nonetheless, it is currently standard practice in road safety modeling to rely on data recorded at the crash scene as reported by the responding police officer to determine such characteristics as who was at-fault in contributing to the crash (Chandraratna & Stamatiadis, 2009; Montella et al., 2013). Crash records in the CARE data used in this study indicate which vehicle (i.e., driver) was deemed to be at-fault in the crash based on the judgement of the police officer responding to the scene and compiling the crash report. While such a designation in the data rests solely on the judgement of one individual, it has been reported elsewhere (e.g., Braitman et al., 2008) that police determination of fault is a reliable measure, certainly given that there are no available alternatives.

Observations with missing or ambiguous values were omitted from the original dataset before performing the model estimation. This yielded a total of 2,008 observed crashes. This study used three injury-severity categories: major injury (fatal or incapacitating injury), minor injury (non-incapacitating injury or possible injury), and property damage only.

Table 2 shows the distribution of the crash factors considered for model estimation. In all, about 4% of the observed crashes were major injury crashes, 13% were minor injury, and 83% were property damage only crashes. Among explanatory variables, CMV drivers were at-fault in 51% of the total crashes, meaning that about 49% of non-CMVs involved in lane change crashes were at-fault. About 66% of the crash observations were sideswipe-same direction and 9% were rear end (front to rear) collisions. Side impact (angled) collisions made up 8% of the crash observations. Further, a gendered analysis shows that 24% of the at-fault drivers were females. About 10% of the at-fault drivers were more than 65 years old. In about 1% of the crashes, the at-fault driver did not buckle up.

4. Results

During model estimation, variables were included in the specification if they had t-statistics corresponding to the 90% confidence interval or above on a two-tailed t-test and the random parameters were included if their standard deviations had t-statistics corresponding to the 90% confidence interval or above. Model estimation results are presented in Table 3, where variables specified for each injury category are denoted by MAJ (major injury), MIN (minor injury), and NO (no injury). In all, 21 crash factors were found to influence the injury severity outcomes of unsafe lane change crashes involving at least one CMV on interstates in Alabama. The McFadden pseudo-ρ^2 value of the model is 0.53, indicates a good fit of the data. According to Chen and Chen (2011), estimated parameters in logit models are not enough to explore the actual influence of a variable on the probability of an injury severity category. Ulfarsson and Mannering (2004) suggests the use of elasticities provided by the average pseudo-elasticity instead of parameter values. The effects of individual factors on injury severity probabilities were therefore explored using elasticities as shown in Table 3. A positive elasticity indicates higher likelihood of a variable to be significantly associated with a particular injury severity and a negative elasticity indicates lower likelihood of a variable to be significantly associated with a particular injury severity.

Four parameter estimates were found to be statistically significant as random parameters: (1) summer; (2) crash location less than 25 miles; (3) four-lane interstate indicator; and (4) CMV driver at-fault. All of the random parameters were shown to be

Table 2
Proportions of crash-related variables used in model estimation.

Variables	Description	Number of observations	Percentage
Dependent			
Major injury	Crash severity: fatal or incapacitating injury	81	4%
Minor injury	Crash severity: non incapacitating or possible injury	261	13%
No injury	Crash severity: property damage only	1,666	83%
Explanatory			
Temporal characteristic			
Summer	Time of the year: Summer (1 = Yes, 0 = No)	563	28%
Environmental characteristic			
Dark	Lighting condition at time of crash: Dark/Unlit (1 = Yes, 0 = No)	321	16%
Location characteristics			
Urban	Crash location: Urban (1 = Yes, 0 = No)	823	41%
Close to home	At-fault driver residence from crash location: <25mi (1 = Yes, 0 = No)	643	32%
Contributing circumstances			
No seatbelt	At-fault driver seatbelt use: No seatbelt (1 = Yes, 0 = No)	21	1%
Blind spot	Contributing circumstance: At-fault driver in blind spot (1 = Yes, 0 = No)	181	9%
Speed	Causal vehicle estimated speed prior to crash: Greater than 70 mph (1 = Yes, 0 = No)	462	23%
Manner of crash			
Sideswipe	Manner of crash: Sideswipe-same direction (1 = Yes, 0 = No)	1,326	66%
Rear end	Manner of crash: rear end (front to rear) (1 = Yes, 0 = No)	181	9%
Side impact	Manner of crash: side impact (angled) (1 = Yes, 0 = No)	161	8%
Roadway characteristics			
Six-lane	Number of lanes: Six-lane interstate (1 = Yes, 0 = No)	884	44%
Four-lane	Number of lanes: Four-lane interstate (1 = Yes, 0 = No)	964	48%
Cable barrier	Interstate median type: Cable barrier (1 = Yes, 0 = No)	121	6%
Straight slope	Geometry: Straight with downward grade (1 = Yes, 0 = No)	241	12%
Driver characteristics			
African American	At-fault driver race: African American (1 = Yes, 0 = No)	643	32%
Female	At-fault driver gender: Female (1 = Yes, 0 = No)	482	24%
Older driver	At-fault driver age: more than 65 (1 = Yes, 0 = No)	201	10%
Driver not impaired	At-fault driver condition at time of crash: Not DUI (1 = Yes, 0 = No)	1,948	97%
Vehicle characteristics			
Major damage	Causal vehicle damage: Major but not disabled (1 = Yes, 0 = No)	422	21%
Trailer	Causal type: Semi-tractor/Semi-trailer (1 = Yes, 0 = No)	743	37%
CMV	Commercial Motor Vehicle at fault (1 = Yes, 0 = No)	1,024	51%

Table 3
Injury-severity model estimation results and elasticities.

Variables	Coefficient	t-Statistic	Elasticity		
			MAJ	MIN	NO
Constant [NO]	1.92	3.82			
Temporal characteristic					
Summer [MAJ]	−1.69	−0.88	38.4%	−2.1%	−1.9%
Standard deviation for Summer (normally distributed)	3.18	2.02			
Location characteristics					
Urban [MAJ]	−0.73	−2.17	−25.3%	1.1%	0.8%
Close to home [MIN]	−0.66	−0.73	−3.4%	23.8%	−4.1%
Standard deviation for Close to home (normally distributed)	2.89	2.45			
Contributing circumstances					
Speed [MAJ]	0.68	1.83	12.7%	−0.7%	−0.6%
No seatbelt [MAJ]	2.92	4.09	1.5%	−0.9%	−0.9%
Blind spot [NO]	−0.68	−2.19	3.5%	3.4%	−0.9%
Manner of crash					
Sideswipe [MAJ]	−1.38	−4.63	−76.6%	3.4%	1.7%
Side impact [MIN]	1.47	4.12	−2.7%	6.0%	0.4%
Rear end [MIN]	1.10	3.48	−12.3%	8.9%	−1.1%
Environmental characteristic					
Dark [MAJ]	0.75	2.24	9.8%	−1.1%	−0.6%
Roadway characteristics					
Straight slope [MAJ]	0.58	1.88	5.6%	−0.6%	−0.3%
Six-lane [MIN]	0.63	2.27	−4.1%	18.1%	−2.8%
Four-lane [NO]	−0.09	−0.17	30.4%	28.6%	−4.8%
Standard deviation for Four-lane (normally distributed)	1.57	1.95			
Cable barrier [NO]	0.93	1.95	−3.7%	−3.6%	0.4%
Vehicle characteristics					
Trailer [MAJ]	−1.22	−3.18	−37.6%	2.0%	0.6%
Major damage [MIN]	−1.03	−3.38	2.1%	−15.4%	1.1%
CMV [NO]	−0.13	−0.27	27.3%	28.8%	−4.8%
Standard deviation for CMV (normally distributed)	1.37	1.75			
Driver characteristics					
Older driver [MIN]	−0.84	−2.07	0.9%	−6.0%	0.4%
Driver not impaired [NO]	1.32	2.93	−81.9%	−76.0%	14.4%
African American [NO]	−0.37	−1.85	6.8%	6.1%	−1.5%
Female [NO]	−0.33	−1.91	5.4%	4.4%	−1.0%
Model statistics					
Number of observations	2,008				
Log-likelihood at constants	−2206.013				
Log-likelihood at convergence	−1036.433				
McFadden pseudoρ^2	0.530				

significantly different from zero with at least 90% confidence. The normal distribution was found to provide the best fit of the modeled random parameters.

The summer variable (specified for major injury) was found to be random with a mean of −1.69 and standard deviation of 3.18. Since the summer indicator variable was assumed to be normally distributed, these numbers indicate that for 29.7% of the summer crashes, the probability of major injury is low. Whereas for the remaining 70.3% of the summer crashes the probability of major injury is high. This shows that the chances of major injury are generally high in about two-thirds of lane change crashes that occur during summer months. The variable for crash location less than 25 miles from the residence of the at-fault driver (specified for minor injury) had mean of −0.66 and standard deviation of 2.89. This means that for 59% of crashes that occurred less than 25 miles from the residence of the at-fault driver, the probability of minor injuries is low while the probability of minor injury is high for the remaining 41% of the crashes. Similar interpretations of mean and standard deviation values for the randomly varying parameters for property damage only lane change crashes show that the four-lane indicator variable increases the probability of property damage only in 28.4% and the CMV driver being at-fault increases the probability of property damage in 46.2% of the crashes. The converse then, is that crashes that occur on four-lane facilities and those that involve CMV drivers being at-fault were less likely to result in property damage only (i.e., at least some form of injury occured) in 71.6% and 53.8% of the lane change crashes, respectively.

Interpretation of the elasticities for parameters specified for major injuries reveal that lane change crashes that occurred during the summer were 38.4% more likely to result in major injuries and those that occurred at urban interstate settings were 25.3% less likely to result in major injuries. Additionally, speed was estimated to increase the probability of major injury by 12.7%, but not wearing a seat belt only increased the probability by 1.5%. It should be noted, however, that seatbelt usage may be under or misreported for non-serious crashes and thus might bias these findings (e.g., Cummings, 2002). Nonetheless, the results reported here reinforce the positive impacts of seatbelt usage. Lane change crashes recorded as sideswipe crashes had 76.6% decreased probability of resulting in major injury and there was a 9.8% increased probability of major injury for lane change crashes that occurred on dark portions of interstates. The probability of a major injury crash outcome for a crash involving an at-fault trailer was found to decrease by 37.6%.

Similar interpretations of elasticities for parameters found to significantly affect (i.e., specified for) minor injury and property damage only crash outcomes can be gleaned from Table 3. For example, lane change crashes involving at-fault CMVs had an increased (28.8%) probability of minor injury. These crashes also exhibited a 27.3% increased probability of being major injury, indicating higher likelihood of some form of injury when CMVs are at-fault in lane changing crashes on interstates. Similarly, there was increased probability of some form of injury (i.e., 3.5% and 3.4% increase in likelihood of major and minor injuries, respectively) for lane change crashes that resulted from the at-fault driver driving in the blind spot of the other vehicle. The presence of cable median barrier was found to decrease the probability of major and minor injury outcomes by 3.7% and 3.6%, respectively. Crashes involving at-fault female drivers had increased probability of major and minor injury outcomes. This is an interesting finding when viewed in terms of the fact that female drivers comprised 24% of at-fault drivers in the data.

5. Discussions

The results provide an enhanced understanding of how a range of contributing factors affect the severity outcomes associated with lane change crashes on interstates. That said, there are limitations to this study presented by the very nature of the data on which it is based. All of the parameters used in the model development were obtained from reports that were compiled by police officers responding to the crash scene. As such, some of the key parameters (i.e., potentially contributing factors) are based solely on the judgement and discretion of the individual officer at the time the crash recorded. Much has been written on the accuracy and reliability of police-reported crash data. For example, with respect to seat belt usage, Schiff and Cummings (2004) reported that police typically identify belted occupants correctly, but they often classify unbelted occupants as belted when they were not actually at the time of the crash. Whereas, Grant et al. (1998) reported that seat belt usage as reported by police agreed with information reported in both the ambulance and emergency room in more than 75% of cases. Others have even documented that police reports sometimes inaccurately record injury severity (Farmer, 2003; Brubacher et al., 2019). On the other hand, Lee et al. (2012) reported good agreement between the data reported by police and crash victim accounts from interviews across a range of crash related characteristics. Nonetheless, concerns over underreporting, inaccuracies, and resulting potential biases in police-reported crash data are a known concern (Abay, 2015; Imprialou & Quddus, 2019) and road safety researchers must be cognizant of these when conducting studies and reporting results.

In many cases, the findings of this study confirm similar results reported elsewhere in the literature. For example, our results suggest that lane change crashes that occur in the dark are more likely to result in major injuries as has been reported in previous studies (e.g., Duncan et al., 1998; Khorashadi et al., 2005; Zhu & Srinivasan, 2011; Islam & Hernandez, 2013; Cerwick et al., 2014). In other cases, the results differed slightly from what has been previously reported – both McGwin and Brown (1999) and Di Stefano and Macdonald (2003) reported a stronger relationship between serious injury lane change crashes and older drivers than this study.

The results clearly indicated that four-lane interstates and the presence of a vehicle in a blind spot are both associated with increased probability of major injury crashes. While there is no definite connection determined herein, it could be inferred that some "unobserved" relationship between fewer travel lanes and a higher occurrence of blind spots is reasonable as vehicles would have less overall maneuverability than if there were additional travel lanes available.

6. Conclusions

This study examined relationships between risk factors and crash outcomes for lane change crashes involving CMV on Alabama interstates. The study used mixed logit modeling technique to account for unobserved heterogeneity across injury severity observations. In all, 21 crash factors were identified to influence injury severity outcomes and four of them were found to be random parameters.

Whether lane change crashes are attributable to certain driver types (older drivers), driving behaviors (speeding), or driving contexts (four-lane interstates), the results presented here reaffirm the need to increase awareness of and ability to safely navigate lane changes in the vicinity of large CMVs. Such efforts can be most effective if they target specific populations of both CMV and non-CMV drivers (Blower & Kostyniuk, 2007; Antin et al., 2017). This study can also serve as foundation for future work on which to base studies investigating the effectiveness of technology-oriented solutions to mitigate lane change crashes, whether vehicle-based systems such as blind spot warning or infrastructure-based technologies that provide dynamic warnings during times or locations of heavy mixed volumes and lane changing activity.

Declaration of interest

None.

Acknowledgment

The authors would like to thank the Center for Advanced Public Safety (CAPS) at the University of Alabama for providing the crash data used for this study. The authors would also thank the Alabama Transportation Institute (ATI) for supporting this study.

References

Abay, K. (2015). Investigating the nature and impact of reporting bias in road crash data. *Transportation Research Part A, 71*, 31–45.

Anastasopoulos, A., & Mannering, F. L. (2011). An empirical assessment of fixed and random parameter logit models using crash- and non-crash-specific injury data. *Accident Analysis & Prevention, 43*(3), 1140–1147.

Antin, J. F., Wotring, B., Perez, M., Glaser, D., 2017. Investigating lane change behaviors and difficulties for senior drivers using the SHRP 2 naturalistic database. National Surface Transportation Safety Center for Excellence Report #17-UM-055.

Bhat, C. (2003). Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. *Transportation Research Part B, 37*(1), 837–855.

Behnood, A., & Mannering, F. (2017). Determinants of bicyclist injury severities in bicycle-vehicle crashes: A random parameters approach with heterogeneity in means and variances. *Analytic Methods in Accident Research, 16*, 35–47.

Blower, D. F., & Kostyniuk, L. P. (2007). *Strategies to reduce CMV-involved crashes, fatalities, and in Michigan UMTRI-2007-26*. Ann Arbor, MI: The University of Michigan Transportation Research Institute.

Braitman, K. A., Kirley, B. B., McCartt, A. T., & Chaudhary, N. K. (2008). Crashes of novice teenage drivers: Characteristics and contributing factors. *Journal of Safety Research, 39*(1), 47–54.

Braver, E. R., Preusser, D. F., Williams, A. F., & Weinstein, H. B. (1996). Major types of fatal crashes between large trucks and cars. In *Proceedings of the 40th annual conference of the Association for the Advancement of Automotive Medicine* (pp. 153–169). Des Plaines, IL: Association for the Advancement of Automotive Medicine.

Brubacher, J., Chan, H., & Erdelyi, S. (2019). Injury severity in police collision reports correlates poorly with requirement for hospital admission. *Journal of Transport & Health, 14* 100606.

Carr, B. (1969). A statistical analysis of rural Ontario traffic accidents using induced exposure data. *Accident Analysis and Prevention, 1*(4), 343–357.

Chandraratna, S., & Stamatiadis, N. (2009). Quasi-induced exposure method: Evaluation of not-at-fault assumption. *Accident Analysis and Prevention, 41*, 308–313.

Chang, L. Y., & Mannering, F. (1999). Analysis of injury severity and vehicle occupancy in truck- and non-truck-involved accidents. *Accident Analysis & Prevention, 31*(5), 579–592.

Chen, F., & Chen, S. (2011). Injury severity of truck drivers in single- and multi-vehicle accidents on rural highways. *Accident Analysis & Prevention, 43*, 1677–1688.

Cerwick, D., Gkritza, K., Shaheed, M., & Hans, Z. (2014). A comparison of the mixed logit and latent class methods for crash severity analysis. *Analytic Methods in Accident Research, 3*, 11–27.

Cummings, P. (2002). Association of seat belt use with death: A comparison of estimates based on data from police and estimates based on data from trained crash investigators. *Injury Prevention, 8*, 338–341.

Di Stefano, M., & Macdonald, W. (2003). Assessment of older drivers: Relationships among on-road errors, medical conditions and test outcome. *Journal of Safety Research, 34*, 415–429.

Duncan, C., Khattak, A., & Council, F. (1998). Applying the ordered probit model to injury severity in truck-passenger car rear-end collisions. *Transportation Research Record, 1635*, 63–71.

Farmer, C. (2003). Reliability of police-reported information for determining crash and injury severity. *Traffic Injury Prevention, 4*, 38–44.

Federal Highway Administration (2017). Infrastructure safety practices for Commercial Motor Vehicles Accessed July 29, 2019 https://safety.fhwa.dot.gov/cmv_rtc/ch1.cfm.

Garber, N. J., Chowdhury, M. A., & Kalaputapu, R. (1992). *Accident characteristics of large trucks on highway ramps* (p. 20005). Washington, DC: AAA Foundation for Traffic Safety.

Garcia, C., Viallon, V., Bouaoun, L., & Martin, J. L. (2019). Prediction of responsibility for drivers and riders involved in injury road crashes. *Journal of Safety Research, 70*, 159–167.

Grants, R., Gregor, M., Maio, R., & Huang, S. (1998). The accuracy of medical records and police reports in determining motor vehicle crash characteristics. *Prehospital Emergency Care, 2*(1), 23–28.

Greene, W. H., Hensher, D. A., & Rose, J. (2006). Accounting for heterogeneity in the variance of unobserved effects in mixed logit models. *Transportation Research Part B, 40*, 75–92.

Halton, J. (1960). On the efficiency of evaluating certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik, 2*, 84–90.

Islam, S., & Jones, S. L. (2014). Pedestrian at-fault crashes on rural and urban roadways in Alabama. *Accident Analysis & Prevention, 72*, 267–276.

Imprialou, M., & Quddus, M. (2019). Crash data quality for road safety research: Current state and future directions. *Accident Analysis and Prevention, 130*, 84–90.

Insurance Institute for Highway Safety and Highway Loss Data Institute. Large Trucks: Fatality Facts. http://www.iihs.org/iihs/topics/t/large-trucks/fatality facts/large-trucks. Accessed December 10, 2018.

Islam, M., & Hernandez, S. (2013). Large truck-involved crashes: Exploratory injury severity analysis. *Journal of Transportation Enginneering, 139*, 596–604.

Islam, S., Jones, S. L., & Dye, D. (2014). Comprehensive analysis of single- and multi-vehicle large truck at-fault crashes on rural and urban roadways in Alabama. *Accident Analysis and Prevention, 67*, 148–158.

Khorashadi, A., Niemeier, D., Shankar, V., & Mannering, F. (2005). Differences in rural and urban driver-injury severities in accidents involving large trucks: An exploratory analysis. *Accident Analysis & Prevention, 37*(5), 910–921.

Kim, J., Ulfarsson, G. F., Kim, S., & Shankar, V. N. (2013). Driver-injury severity in single-vehicle crashes in California: A mixed-logit analysis of heterogeneity due to age and gender. *Accident Analysis and Prevention., 50*, 1073–1081.

Lee, Y. M., Platts-Mills, T. F., MacWilliams, J. B., Sochor, M. R., Jones, J. S., Domeier, R. M., Schneider, L. W., & McLean, S. A. (2012). Descriptions of motor vehicle collisions by participants in emergency department–based studies: are they accurate? *Western journal of emergency medicine, 13*(4), 329.

Mannering, F., & Bhat, C. (2014). Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research, 1*, 1–22.

Mannering, F., Shankar, V., & Bhat, C. (2016). Unobserved heterogeneity and the statistical analysis of highway accident data. *Analytic Methods in Accident Research, 11*, 1–16.

McFadden, D. (1981). Econometric models of probabilistic choice. In C. F. Manski & D. McFadden (Eds.), *A structural analysis of discrete data with econometric applications*. Cambridge, MA: The MIT Press.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics, 15*(5), 447–470.

McGwin, G., Jr., & Brown, D. B. (1999). Characteristics of traffic crashes among young, middle-aged, and older drivers. *Accident Analysis and Preventions, 31*, 181–198.

Milton, J., Shankar, V., & Mannering, F. (2008). Highway accident severities and the mixed logit model: An exploratory empirical analysis. *Accident Analysis & Prevention, 40*(1), 260–266.

Montella, A., Andreassen, D., Tarko, A., Turner, S., Mauriello, F., Imbriani, L., & Romero, M. (2013). Crash Databases in Australasia, the European Union, and the United States: Review and prospects for improvement. *Transportation Research Record, 2386*(1), 128–136.

Morgan, A., & Mannering, F. (2011). The effects of road-surface conditions, age, and gender on driver-injury severities. *Accident Analysis & Prevention, 43*(5), 1852–1863.

National Highway Traffic Safety Administration, 2017. Fatality analysis reporting system (FARS). Available online at http://www-fars.nhtsa.dot.gov.

Pahukula, J., Hernandez, S., & Unnikrishnan, A. (2015). A time of day analysis of crashes involving large trucks in urban areas. *Accident; Analysis and Prevention, 75*, 155–163.

Roberston, M., & Drummer, O. (1994). Responsibility analysis: A methodology to study the effects of drugs in driving. *Accident Analysis and Prevention, 26*(2), 243–247.

Salmi, L., Orriols, L., & Lagarde, E. (2012). Comparing responsible and non-responsible drivers to assess determinants of road traffic collisions: Time to standardise and revisit. *Injury Prevention, 20*(6), 380–386.

Savolainen, P., Mannering, F., Lord, D., & Quddus, M. (2011). The statistical analysis of crash-injury severities: A review and assessment of methodological alternatives. *Accident Analysis and Prevention, 43*(5), 1666–1676.

Schiff, M., & Cummings, P. (2004). Comparison of reporting of seat belt use by police and crash investigators: Variation in agreement by injury severity. *Accident Analysis and Prevention, 36*(6), 961–965.

Sen B., Smith J. D., Najm W. G, 2003. Analysis of lane change crashes, Performed by John A. Volpe National Transportation System Center, Cambridge, MA, Sponsored by National Highway Traffic Safety Administration, Washington DC, DOT HS 809 571.

Seraneeprakarn, P., Huang, S., Shankar, V., Mannering, F., Venkataraman, N., & Milton, J. (2017). Occupant injury severities in hybrid-vehicle involved crashes: A random parameters approach with heterogeneity in means and variances. *Analytic Methods in Accident Research, 15*, 41–55.

Shaheed, M. S. B., Gkritza, K., Zhang, W., & Hans, Z. (2013). A mixed logit analysis of two-vehicle crash severities involving a motorcycle. *Accident Analysis and Prevention, 61*, 119–128.

Train, K. (2003). *Discrete choice methods with simulation*. Cambridge, U.K.: Cambridge University Press.

Train, K. 1999. Halton sequences for mixed logit. Working Paper, Department of Economics, University of California, Berkley.

U.S. Department of Transportation. Bureau of Transportation Statistics. Commodity Flow Surver Overview https://www.bts.gov/cfs Accessed March 11, 2019.

U.S. Department of Transportation-Federal Motor Carrier Safety Administration Analysis Division 2018. Understanding Truck and Bus challenges – Our roads, our safety. Available online at https://www.fmcsa.dot.gov/ourroads/cmv-safety-challenges.

Ulfarsson, G., & Mannering, F. (2004). Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. *Accident Analysis and Prevention, 36*(2), 135–147.

Venkataraman, N., Ulfarsson, G., & Shankar, V. (2013). Random parameter models of interstate crash frequencies by severity, number of vehicles involved, collision and location type. *Accident Analysis and Prevention, 59*, 309–318.

Venkataraman, N., Shankar, V., Ulfarsson, G., & Deptuch, D. (2014). A heterogeneity-in-means count model for evaluating the effects of interchange type on heterogeneous influences of interstate geometrics on crash frequencies. *Analytic Methods in Accident Research, 2*, 12–20.

Wang, H.L. & Shi, Z.K. (2013). Hierarchical regression model for truck collision severity analysis. In Advanced Materials Research (Vol. 671, pp. 2889–2892). Trans Tech Publications Ltd.

Washington, S. P., Karlaftis, M. G., Mannering, F. L., 2011. Statistical and econometric methods for transportation data analysis, second ed. Chapman &Hall/CRC.

Zhu, H., & Srinivasan, S. (2011). A comprehensive analysis of factors influencing the injury severity of large-truck crashes. *Accident analysis and Prevention, 43*, 49–57.

Emmanuel Kofi Adanu is an Associate Research Engineer at the Alabama Transportation Institute at the University of Alabama.

Abhay Lidbe is an Associate Research Engineer at the Alabama Transportation Institute at the University of Alabama.

Elsa Tedla is an assistant research engineer at the Alabama Transportation Institute at the University of Alabama.

Steven Jones is a professor at the Department of Civil, Construction, and Environmental Engineering and the Deputy Director of Alabama Transportation Institution at the University of Alabama.

Journal of Safety Research 76 (2021) 248–255

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Investigating hazardous factors affecting freeway crash injury severity incorporating real-time weather data: Using a Bayesian multinomial logit model with conditional autoregressive priors

Xuan Zhang [a], Huiying Wen [a,*], Toshiyuki Yamamoto [b], Qiang Zeng [a]

[a] School of Civil Engineering and Transportation, South China University of Technology, Guangzhou, Guangdong, 510641, PR China
[b] Institute of Materials and Systems for Sustainability, Nagoya University, Nagoya 464-8603, Japan

ARTICLE INFO

Article history:
Received 9 April 2020
Received in revised form 22 September 2020
Accepted 16 December 2020
Available online 7 January 2021

Keywords:
Freeway safety
Crash injury severity
Real-time weather data
Multinomial logit model
Spatial effect

ABSTRACT

Introduction: It has been demonstrated that weather conditions have significant impacts on freeway safety. However, when employing an econometric model to examine freeway crash injury severity, most of the existing studies tend to categorize several different adverse weather conditions such as rainy, snowy, and windy conditions into one category, "adverse weather," which might lead to a large amount of information loss and estimation bias. Hence, to overcome this issue, real-time weather data, the value of meteorological elements when crashes occurred, are incorporated into the dataset for freeway crash injury analysis in this study. *Methods:* Due to the possible existence of spatial correlations in freeway crash injury data, this study presents a new method, the spatial multinomial logit (SMNL) model, to consider the spatial effects in the framework of the multinomial logit (MNL) model. In the SMNL model, the Gaussian conditional autoregressive (CAR) prior is adopted to capture the spatial correlation. In this study, the model results of the SMNL model are compared with the model results of the traditional multinomial logit (MNL) model. In addition, Bayesian inference is adopted to estimate the parameters of these two models. *Result:* The result of the SMNL model shows the significance of the spatial terms, which demonstrates the existence of spatial correlation. In addition, the SMNL model has a better model fitting ability than the MNL model. Through the parameter estimate results, risk factors such as vertical grade, visibility, emergency medical services (EMS) response time, and vehicle type have significant effects on freeway injury severity. *Practical Application:* According to the results, corresponding countermeasures for freeway roadway design, traffic management, and vehicle design are proposed to improve freeway safety. For example, steep slopes should be avoided if possible, and in-lane rumble strips should be recommended for steep down-slope segments. Besides, traffic volume proportion of large vehicles should be limited when the wind speed exceeds a certain grade.

1. Introduction

Freeways, with large traffic volumes and high vehicle speeds, play an important role in the comprehensive system of transportation. With the increase in freeway mileage worldwide, freeway safety problems are becoming much more challenging than before. According to the statistics from the Traffic Management Bureau of the Public Security Ministry in China, 2010–2015, high fatality and injury rates have been the key features of freeway crashes. Currently, for the purpose of determining the freeway crash mechanism and reducing the fatality and injury rate, it is popular to establish an econometric model to explore the link between freeway crash injury severity and risk factors (including human factors, traffic flow characteristics, roadway factors, vehicle characteristics, and environmental characteristics) (Golob et al., 1987; Haleem & Gan, 2013; Ma et al., 2016; Mergia et al., 2013; Shankar et al., 1996; Yu & Abdel-Aty, 2014; Zhang et al., 2011). After identifying the hazardous factors, corresponding countermeasures for engineering and management are proposed.

Weather conditions, an important part of the environmental characteristics, have been demonstrated to significantly influence freeway crash injury severity. For instance, through the data collected in Florida, United States, Zhang et al. (2011) found that when the weather is clear, the probability of no injury on freeway diverging areas increases, and the probability for all levels of injury

* Corresponding author.
E-mail addresses: zhangxuanscut@163.com (X. Zhang), hywen@scut.edu.cn (H. Wen), yamamoto@civil.nagoya-u.ac.jp (T. Yamamoto), zengqiang@scut.edu.cn (Q. Zeng).

https://doi.org/10.1016/j.jsr.2020.12.014
0022-4375/

decreases. With the police-recorded crash data in Ohio, Mergia et al. (2013) demonstrated that adverse weather is accompanied by a higher risk for fatality when the crash occurs in the freeway merging area. However, in the existing freeway crash severity analysis, the definition of weather condition indicator has some limitations. For example, Mergia et al. (2013) defined foggy, rainy, snowy, windy, and other weather conditions as "adverse weather conditions." Zhang et al. (2011) separated weather conditions into two categories, "clear day" and "other." The category "other" includes rainy, snowy, and foggy conditions. It is well known that weather conditions such as windy, rainy, and foggy conditions are totally different from each other and have distinct impacts on drivers' driving behavior and vehicle operation. For example, on rainy days, the roadway is wet, and the visibility is low, which adversely impacts the friction of the vehicle tires and the driver's field of vision. In comparison, on windy days, it is easy for vehicles to overturn, but the friction of the vehicle tires and the driver's vision are not affected. Therefore, combining these weather conditions into one category in the model might lead to a large amount of information loss and biased results. To overcome this issue, incorporating real-time weather data (the values of weather elements such as wind speed and rainfall amount at the time of the accident) in a crash analysis dataset might be a good alternative, with the advantage of describing weather conditions more accurately and comprehensively. Researchers have evaluated crash injury severity based on real-time weather data in some fields, for example, in pedestrian crashes (Zhai et al., 2019), urban arterial crashes (Theofilatos & Yannis, 2017), single-vehicle crashes (Jung et al., 2010), and truck-related crashes (Naik et al., 2016). To the best of the authors' knowledge, to date, only Yu and Abdel-Aty (2014) has explored the links between freeway crash injury severity and hazardous factors incorporating real-time weather data. However, Yu and Abdel-Aty (2014) only categorized freeway crash injury into two levels: severe crashes and non-severe crashes. The information provided is limited. Hence, in this paper, freeway crash injury severity analysis with real-time weather data incorporated is conducted, and the injury severity is divided into more than two categories.

Regarding methodology, the multinomial logit (MNL) model has been used frequently when analyzing crash injury severity with more than two injury severity levels (Carson & Mannering, 2001; Chen & Fan, 2019; Malyshkina & Mannering, 2010; Shankar & Mannering, 1996; Ulfarsson & Mannering, 2004; Ye & Lord, 2014). However, few studies consider spatial effects when adopting the MNL model to analyze crash severity. To date, it has been a trend to consider spatial effects in crash severity studies (Meng et al., 2017; Xu et al., 2016; Zeng et al., 2019; Zeng, Hao, et al., 2020) because some unobserved factors might commonly affect adjacent segments, resulting in spatial correlations in roadway crash data (Huang et al., 2016; Ma et al., 2017; Zeng et al., 2017; Zeng & Huang, 2014a; Zeng, Wen, et al., 2020). Moreover, for this research, there is another important reason to take into account spatial correlation. As pointed out by Shankar et al. (1995), when using the data observed by weather stations, several continuous freeway segments sharing the data of one common weather station might lead to spatial correlation in the dataset. Ignoring the spatial effect might result in estimate bias. To address this issue, this paper presents a new approach, the spatial multinomial logit (SMNL) model, with the MNL model serving as the basic framework and the Gaussian conditional autoregressive (CAR) prior employed to account for the spatial correlation. It has been demonstrated by several studies (Meng et al., 2017; Xu et al., 2016; Zeng et al., 2019) that the Gaussian CAR prior in the crash severity model could well capture the spatial effect. Besides, with the help of the free software WinBUGS, the CAR priors could be introduced to multinomial logit model easily.

In summary, in this research, real-time weather data are incorporated when examining the relationship between freeway crash injury severity and hazardous factors. To provide more information about injuries, the injury severity is categorized into more than two levels. In addition, a new method, the SMNL model, is proposed and compared with the traditional MNL when fitting freeway crash injury data.

The rest of this article is structured as follows. Section 2 describes the freeway crash analysis dataset. Section 3 outlines the structure of the MNL model and the SMNL model, the estimation process, and the calculation method of the marginal effect. Section 4 presents the estimate results of the SMNL model and the MNL model, the marginal effect of the significant factors and the discussion. The conclusion and future research are shown in Section 5.

2. Data

The freeway crash injury dataset analyzed in this study is from Kaiyang Freeway, a 125.2 km long four-lane double carriageway freeway in Guangdong Province, China. The data of 2014 and 2015 were collected. This dataset originates from three datasets: (1) the freeway crash dataset, (2) freeway road geometry file, and (3) meteorological dataset. The freeway crash dataset is extracted from the Highway Maintenance and Administration Management System, maintained by the Guangdong Transportation Group. The freeway road geometry file is maintained by the Guangdong Province Communication Planning and Design Institute Company Limited. The meteorological dataset is collected from the Meteorological Information Management System maintained by the Guangdong Climate Center.

There are 1,420 crashes reported in the freeway crash dataset. Among them, 1,152 crashes are no injury crashes, 205 crashes are minor injury crashes, 30 crashes are severe injury crashes, and 33 crashes are fatal crashes. Since both the numbers of severe injury crashes and fatal crashes are too small, in this paper, these two injury levels are combined into one injury severity category.

In addition to the crash injury information, the freeway crash dataset includes other detailed information for every crash record as follows: the mileage of the crash location, the specific time when the crash occurred, the collision type, whether the involved driver is professional or not, the type of vehicle involved, the license number, and the emergency medical services (EMS) response time. According to this information, the exogenous variables are arranged, as shown in Table 1 and Table 2, and the detailed explanations are in Table 1 and Table 2.

Concerning providing more information about the roadway geometry of the crash location, we divide the Kaiyang freeway into 154 consecutive segments according to the homogeneity of the horizontal curvature and vertical grade. Then, the road feature dataset is sorted from the freeway road geometry file, including the horizontal curvature value, vertical grade value of each segment, whether the segment is part of a bridge, and whether the segment is close to the ramp. Finally, according to the mileage of the crash location, the corresponding segment is matched, and detailed road feature information is recorded in the road feature dataset.

The meteorological dataset contains every hour of weather factor data including the wind speed, visibility, and rainfall amount from three weather stations adjacent to the Kaiyang Freeway from 2014 to 2015. These three weather stations are the Kaiping station, Enping station, and Yangjiang station. According to the crash location and the specific time that the crash occurred, the meteorological data from the same hour and location are matched. Since precipitation, visibility and wind are the most commonly used

Table 1
Data description for response variable and binary variables.

Subtitle	Variable	Description		count	ratio
Response variable					
	Injury severity level	No injury = 1	1	1,152	81.13%
		Minor injury = 2	2	205	14.44%
		Severe injury & fatality = 3	3	63	4.43%
Explanatory variable (binary variable)					
Crash time	Before dawn*	Crash occurrence time is in the period from 12 a.m. to 6 a.m. = 1; otherwise = 0	1	260	18.31%
			0	1,160	81.69%
	Morning	Crash occurrence time is in the period from 6 a.m. to 12 p.m. = 1; otherwise = 0	1	315	22.18%
			0	1,105	77.82%
	Afternoon	Crash occurrence time is in the period from 12 p.m. to 6 p.m. = 1; otherwise = 0	1	529	37.25%
			0	891	62.75%
	Evening	Crash occurrence time is in the period from 6 p.m. to 12 a.m. = 1; otherwise = 0	1	316	22.25%
			0	1,104	77.75%
Vehicle type	Coach	At least one involved vehicle is a coach = 1; otherwise = 0	1	91	6.41%
			0	1,329	93.59%
	Truck	At least one involved vehicle is a truck = 1; otherwise = 0	1	447	31.48%
			0	973	68.52%
	Other vehicle	At least one other vehicle (such as a trailer) is involved = 1; otherwise = 0	1	142	10.00%
			0	1,278	90.00%
Roadway feature	Bridge	The crash location is on a bridge = 1; otherwise = 0	1	762	53.66%
			0	658	46.34%
	Ramp	The crash location is near a ramp = 1; otherwise = 0	1	347	24.44%
			0	1,073	75.56%
Other Binary Variables	Nonlocal vehicle	All the involved vehicles are registered in Guangdong Province = 0; otherwise	1	142	10.00%
		(at least one involved vehicle is a non-local vehicle) = 1	0	1,278	90.00%
	Driver type	None of the involved drivers are professional drivers = 0; otherwise = 1	1	55	3.87%
			0	1,365	96.13%
	Single-vehicle crash	Only one car is involved in the crash = 1; otherwise = 0	1	650	45.77%
			0	770	54.23%

* The reference category.

Table 2
Data description for continuous variables.

Explanatory variable	Description	Mean	S.D.	Max.	Min.
EMS response time	Duration between crash report time and the arrival time of EMS (min)	19.38	16.65	260	0
Horizontal curvature	The horizontal curvature of the freeway segment where the crash occurred (0.1 km^{-1})	1.84	1.23	4.35	0
Vertical grade	The vertical grade of the freeway segment where the crash occurred (%)	0.71	0.59	2.91	0
Max wind speed	The maximum wind speed during the hour when the crash occurred (m/s)	3.83	2.06	17.3	0.7
Rainfall amount	The total rainfall amount during the hour when the crash occurred (mm)	0.76	3.43	54.8	0
Min visibility	The minimum visibility during the hour when the crash occurred (km)	18.10	18.73	80	0.1

weather characteristics when analyzing crash data (Theofilatos & Yannis, 2014), in this paper, the maximum wind speed value, minimum visibility value, and rainfall amount value during the hour that a crash occurred serve as explanatory variables in the models. To avoid the strong correlations among meteorological elements, Pearson correlation test is conducted and the result shows that the correlation among the maximum wind speed value, minimum visibility value, and rainfall amount value are weak (the coefficients are below 0.3 or over −0.3).

The data description of the freeway crash dataset is shown in Table 1 and Table 2. The data description of the response variable and binary explanatory variables is presented in Table 1. The data description of the continuous variables is displayed in Table 2.

3. Methodology

In this section, first, the model structures of the multinomial logit (MNL) model and spatial multinomial logit (SMNL) model are presented. Then, the estimation processes of the MNL model and SMNL model, the comparison criteria for model fitting, and the method to calculate how the variation in the significant factors impacts the probability of each injury severity level would be shown.

3.1. Model specifications

3.1.1. Multinomial logit model

As mentioned in Section 2, the crash injury severity in this study is categorized into three levels: 1 (no injury), 2 (minor injury), and 3 (severe injury and fatality). For the three severity levels, the probability of crash i with injury severity n can be written as:

$$P_{i,n} = P(U_{i,n} \geqslant U_{i,m}), \forall m \neq n \tag{1}$$

where $P_{i,n}$ is the probability of crash i with injury level n, P denotes the probability and U_{in} serves as the utility function that determines the injury outcome of crash i to be severity level n. The utility function can be written as:

$$U_{i,n} = \boldsymbol{\beta_n X_i} + \varepsilon_{in} \tag{2}$$

where $\boldsymbol{X_i}$ represents a vector of explanatory variables that determine the crash injury severity, $\boldsymbol{\beta_n}$ represents the corresponding coefficient for injury severity n, and ε_{in} is the error term that accounts for the unobserved factors affecting the crash injury severity level. ε_{in} is assumed to be independently distributed. Based on Eqs. (1) and (2), the following equation can be written as:

$$P_{i,n} = P(\boldsymbol{\beta_n X_i} - \boldsymbol{\beta_m X_i} \geqslant \varepsilon_{im} - \varepsilon_{in}), \forall m \neq n \tag{3}$$

In Eq. (3), the error term ε_{in} is assumed to be a generalized extreme value (GEV) distribution. Finally, the probability for crash i with injury severity outcome $n$can be calculated as follows:

$$P_{i,n} = \exp(\boldsymbol{\beta_n X}_i) / \sum_{m=1}^{M} \exp(\boldsymbol{\beta_m X}_i) \tag{4}$$

Specifically, Eq. (4) can be rewritten as Eq. (5), with injury severity 1 serving as the reference category for the other injury severity levels, and M indicates the total number of injury severity levels:

$$\begin{aligned} P_{i,n} &= \exp(\boldsymbol{\beta_n X}_i) / \sum_{m=1}^{M} \exp(\boldsymbol{\beta_m X}_i) \\ &= [\exp(\boldsymbol{\beta_n X}_i)/\exp(\boldsymbol{\beta_1 X}_i)] / \left\{ \sum_{m=1}^{M} [\exp(\boldsymbol{\beta_m X}_i)/\exp(\boldsymbol{\beta_1 X}_i)] \right\} \\ &= \exp[(\boldsymbol{\beta_n} - \boldsymbol{\beta_1})\boldsymbol{X}_i] / \left\{ 1 + \sum_{m=2}^{M} \exp[(\boldsymbol{\beta_m} - \boldsymbol{\beta_1})\boldsymbol{X}_i)] \right\} \\ &= \begin{cases} 1/[1 + \sum_{m=2}^{M} \exp(\boldsymbol{\alpha_m X}_i)], n = 1 \\ \exp(\boldsymbol{\alpha_n X}_i)/[1 + \sum_{m=2}^{M} \exp(\boldsymbol{\alpha_m X}_i)], n \neq 1 \end{cases} \end{aligned} \tag{5}$$

where $\boldsymbol{\alpha_n}$ serves as the coefficient parameters to be estimated instead of the original parameters $\boldsymbol{\beta_n}$.

3.1.2. Spatial multinomial logit model

As mentioned in section 1, the Gaussian CAR prior can capture the spatial correlation well when evaluating crash severity (Meng et al., 2017; Xu et al., 2016; Zeng et al., 2019). Specifically, in this study, for a crash i occurring on the segment k with injury severity level n, the utility function can be modified to:

$$U_{in} = \boldsymbol{\beta_n X}_i + \phi_{nk} + \varepsilon_{in} \tag{6}$$

where the term ϕ_{nk} denotes the spatial correlation of each crash on freeway segment k for injury severity n and is assumed to follow the CAR Gaussian distribution:

$$\phi_{nk} \sim N\left(\frac{\sum_{k \neq j} \phi_{nj} \omega_{kj}}{\sum_{k \neq j} \omega_{kj}}, \frac{\sigma_{\phi_n}}{\sum_{k \neq j} \omega_{kj}}\right) \tag{7}$$

where ω_{kj} is the proximity weight between the freeway segments j and k. The proximity weight is defined by the binary first-order proximity structure, which has been widely used in previous research (Meng et al., 2017; Xu et al., 2016; Zeng et al., 2019). Specifically, if segments j and k are connected, ω_{kj}=1; otherwise, ω_{kj}=0. σ_{ϕ_n} is the variance parameter of the spatial term for each injury severity level.

Similar to Eq. (5), the probability of crash i with injury outcome n is calculated as follows (with injury severity level 1 serving as the reference category):

$$\begin{aligned} P_{i,n} &= \exp(\boldsymbol{\beta_n X}_i + \phi_{nk}) / \sum_{m=1}^{M} \exp(\boldsymbol{\beta_m X}_i + \phi_{mk}) \\ &= [\exp(\boldsymbol{\beta_n X}_i + \phi_{nk})/\exp(\boldsymbol{\beta_1 X}_i + \phi_{1k})] / \left\{ \sum_{m=1}^{M} [\exp(\boldsymbol{\beta_m X}_i + \phi_{mk})/\exp(\boldsymbol{\beta_1 X}_i + \phi_{1k})] \right\} \\ &= \exp[(\boldsymbol{\beta_n} - \boldsymbol{\beta_1})\boldsymbol{X}_i + \phi_{nk} - \phi_{1k}] / \left\{ 1 + \sum_{m=2}^{M} \exp[(\boldsymbol{\beta_m} - \boldsymbol{\beta_1})\boldsymbol{X}_i + \phi_{mk} - \phi_{1k}] \right\} \\ &= \begin{cases} 1/[1 + \sum_{m=2}^{M} \exp(\boldsymbol{\alpha_m X}_i + \phi_{mk} - \phi_{1k})], \ n = 1 \\ \exp(\boldsymbol{\alpha_n X}_i + \phi_{nk} - \phi_{1k})/[1 + \sum_{m=2}^{M} \exp(\boldsymbol{\alpha_m X}_i + \phi_{mk} - \phi_{1k})], n \neq 1 \end{cases} \end{aligned} \tag{8}$$

3.2. Model estimation

Because the maximum likelihood estimation cannot be used to estimate models with Gaussian CAR priors, in this study, the Bayesian method is adopted to estimate the parameters of the model. The Bayesian method is based on Markov Chain Monte Carlo (MCMC) simulation with the Gibbs sampling algorithm, and through the free software WinBUGS, this method can be realized easily. To employ the Bayesian method to estimate the parameters, first, the prior distribution of each parameter and hyperparameter should be specified in the model. In this study, due to the lack of additional knowledge, a noninformative prior distribution is used for the parameters $\boldsymbol{\alpha_n}$ and hyperparameters σ_{ϕ_n}. Specifically, a diffused normal distribution, $N(0, 10^4)$, is used as the priors of $\boldsymbol{\alpha_n}$, the coefficient parameters. A diffuse gamma distribution denoted by $gamma(0.001, 0.001)$ for the priors of the precision parameters, $1/\sigma_{\phi_n}$ would be adopted. For each model, a chain of 200,000 MCMC simulation iterations is run, where the first 150,000 iterations serve as the burn-in. To ensure that the simulation converges, the MCMC trace plot for the parameters is checked.

3.3. Model comparison criterion

In this paper, deviance information criterion (DIC) and the classification accuracy for the entire dataset and for each injury severity n are used to compare the model fitting ability.

According to Spiegelhalter et al. (2002), the definition of DIC is shown as Eq. (9):

$$\text{DIC} = \bar{D} + pD \tag{9}$$

$\bar{D}$ is the posterior mean deviance used to measure the fitness or adequacy of the model, and a lower $\bar{D}$ indicates better model fitting. pD is the effective number of parameters, which is used to measure the complexity of the model. Most of the time, the model is preferred if it has a lower DIC value. The values of $\bar{D}$, pD and DIC can be obtained directly from the software WinBUGS.

The classification accuracy for the entire dataset, defined as the proportion of accurate prediction in the entire dataset, is calculated as (Zeng et al., 2019; Zeng & Huang, 2014b):

$$\text{CA}_\text{t} = \frac{\sum_{\bar{Y}_i = Y_i} Y_i/Y_i}{\sum_i Y_i/Y_i} \times 100\% \tag{10}$$

where Y_i represents the predicted crash severity level.

In addition, the classification accuracy for injury severity level n, defined as the proportion of accurate predictions in the dataset with observed injury severity level n, is calculated as:

$$\text{CA}_\text{n} = \frac{\sum_{\bar{Y}_i = Y_i = n} Y_i/Y_i}{\sum_{Y_i = n} Y_i/Y_i} \times 100\% \tag{11}$$

3.4. The calculation process of the marginal effect

Through the parameter estimate results of the MNL model and SMNL model, it is hard to see how the variation in the contributing factors directly influences the probability of each injury severity level. Therefore, the marginal effect is used to analyze this impact in this study. Since the explanatory variable in this paper contains continuous variables and binary indicator variables, the calculation of the marginal effect for these two kinds of variables is shown separately.

For continuous variables, the calculation of the marginal effect for each crash injury severity n is calculated as follows (Scott Long, 1997):

$$\frac{\partial P_{i,n}}{\partial x} = \begin{cases} P_{i,1}(-\sum_{m=2}^{M} \alpha_m^x P_{i,m}), n = 1 \\ P_{i,n}(\alpha_n^x - \sum_{m=2}^{M} \alpha_m^x P_{i,m}), n \neq 1 \end{cases} \quad (12)$$

where α_n^x is the corresponding coefficient for the specific variable x in the probability function of crash injury severity level n.

For binary indicator variables, the calculation of the marginal effect for each crash injury severity n is shown in Eq. (13) (Scott Long, 1997):

$$\frac{\Delta P_{i,n}}{\Delta x} = P_{i,n}(x = 1) - P_{i,n}(x = 0) \quad (13)$$

4. Results

4.1. Model comparison

First, through Table 3, it can easily be found that the standard deviation of the spatial terms for the no injury level and minor injury level, $sd(\phi_1)$ and $sd(\phi_2)$, are significant in the SMNL model, suggesting the existence of spatial correlation. The spatial effect might be attributed to several continuous freeway segments sharing the data of one common weather station and the unobserved factors that clustered spatially such as the lighting conditions and terrain features.

Second, the fitting ability of these two models is compared in this section. Comparison of the DIC values shows that the MNL model has a lower DIC value (DIC value for MNL model = 1552.84) than the SMNL model (DIC value for SMNL model = 1578.16). However, the SMNL model has a much lower $\bar{D}$ value and a higher value of pD than the MNL model because the SMNL model contains spatial terms for three injury levels and thus has higher model complexity. The lower $\bar{D}$ value of the SMNL model suggests that the goodness-of-fit of the SMNL model is better than that of the MNL model. For the classification accuracy for the entire dataset, it can be easily found that the SMNL model has higher classification accuracy (82.39%) than the MNL model (80.99%). Additionally, through the results of the classification accuracy for each injury severity level, although having a lower classification accuracy for no injury level, the SMNL model has higher classification accuracy for the minor injury level (CA_2=15.12%) and severe injury & fatality level (CA_3=14.29%) than the MNL model (CA_2=2.44%,CA_3 = 4.76%). On the whole, although having a higher DIC value, the SMNL model has better model fitting ability (lower $\bar{D}$ value), higher classification accuracy for the entire dataset and higher classification accuracy for the minor injury level and severe injury & fatality level (which is important when reducing the injury and fatality rate), indicating that the SMNL model outperforms the MNL model.

Moreover, from Table 3, the estimate results suggest that the coefficients are generally consistent between the MNL model and the SMNL model. Therefore, in this paper, the result of the SMNL model is chosen as the main model to illustrate how the contributory factors influence the freeway crash injury severity level.

4.2. Parameter estimates

As mentioned above, the result of the SMNL model is selected as the main model to discuss how the contributory factors influence freeway crash severity. Additionally, with the purpose of showing the impacts directly, in this section, the marginal effects of the significant factors in the MNL model and SMNL model are calculated and shown in Table 4.

The results of the SMNL model show that if the vertical grade of crash location is enhanced by 1%, the probabilities of *no injury* and *minor injury* decrease by 2.09% and 0.77%, respectively. In addition, the probability of *severe injury and fatality* increases by 2.86%. This result indicates that a steep freeway segment is accompanied by a worse crash injury severity. This result is consistent with previous studies. As pointed out by previous studies (Christoforou et al., 2010; Savolainen & Mannering, 2007; Yu & Abdel-Aty, 2014), a possible reason for this result is that a steeper segment causes a shorter sight distance, leading to less time for drivers to respond when a hazardous scene occurs. In addition, Lu et al. (2016) mentioned that a steep slope might result in two risky situations, increasing vehicle speed and heating of pavement.

Table 3
Parameter estimation and model comparison result.

Model	MNL model		SMNL model	
	Injury level 2	Injury level 3	Injury level 2	Injury level 3
Afternoon	−0.43 (0.25)*	−1 (0.48)**	–	−1.08 (0.51)**
Coach	0.6 (0.27)**	–	0.69 (0.3)**	–
Truck	0.54 (0.18)**	1.1 (0.31)**	0.56 (0.2)**	1.17 (0.33)**
Other vehicle	0.45 (0.25)*	–	–	–
Nonlocal vehicle	0.34 (0.18)*	-	0.33 (0.2)*	–
EMS response time	0.01 (0.005)**	0.02 (0.008)**	0.01 (0.006)*	0.02 (0.009)**
Max wind speed	–	−0.24 (0.1)**	–	–
Min visibility	0.008 (0.005)*	–	0.01 (0.005)**	–
Vertical grade	–	0.58 (0.2)**	–	0.71 (0.3)**
Single-vehicle crash	−1.04 (0.19)**	−1.09 (0.35)**	−1.19 (0.21)**	−1.23 (0.37)**
constant	−1.46 (0.36)**	−3.04 (0.62)**	−9.44 (1.3)**	−17.81 (1.60)**
$sd(\phi_1)$	–		0.33 (0.31)**	
$sd(\phi_2)$	–		0.95 (0.46)**	
$sd(\phi_3)$	–		0.49 (0.53)	
$\bar{D}$	1517.77		1455.83	
pD	35.064		122.325	
DIC	1552.84		1578.16	
CA_1	99.13%		98.09%	
CA_2	2.44%		15.12%	
CA_3	4.76%		14.29%	
CA_t	80.99%		82.39%	

* Indicates significance at the 90% credibility level.
** Indicates significance at the 95% credibility level.

Table 4
Marginal effects for the MNL and SMNL models.

	MNL model			SMNL model		
	No injury (%)	Minor injury (%)	Severe injury & fatality (%)	No injury (%)	Minor injury (%)	Severe injury & fatality (%)
Continuous variable						
Vertical grade	−1.7	−0.6	2.3	−2.09	−0.77	2.86
Min visibility	−0.1	0.1	0	−0.1	0.1	0
EMS response time	−0.16	0.1	0.06	−0.17	0.1	0.07
Max wind speed	0.7	0.2	−0.9	0	0	0
Binary variable						
Coach	−7.51	8.16	−0.65	−8.72	9.54	−0.82
Truck	−9.61	5.58	4.03	−10.1	5.8	4.3
Other vehicle	−5.40	5.90	−0.50	0	0	0
Nonlocal vehicle	−3.83	4.16	−0.33	−3.65	4	−0.35
Afternoon	7.11	−4.08	−3.03	2.7	0.9	−3.6
Single-vehicle crash	14.09	−10.84	−3.25	16.11	−12.42	−3.69

Visibility is shown to have a significant effect on freeway crash injury severity. In the results of the SMNL model, a one-unit decrease in *minimum visibility* leads to a 0.1% increase in the probability of *no injury* and a 0.1% decrease in the probability of *minor injury*, indicating that low visibility leads to less severe injury outcomes. This result is consistent with previous results (Hou et al., 2019). A possible reason might be drivers' risk compensation. When visibility decreases, drivers tend to be more cautious and start to decrease speed. The gain in drivers' alertness in low visibility conditions has been demonstrated by some previous studies (Jeihani & Banerjee, 2018; Mueller & Trick, 2012) through conducting driving simulator experiments.

EMS response time is shown to significantly affect freeway crash injury severity. With a one-minute increase in *EMS response time*, the probabilities for *minor injury* and *severe injury and fatality* increase by 0.1% and 0.07%, respectively, while the probability for *no injury* decreases by 0.17%. This result is consistent with previous studies (Gonzalez et al., 2009; Lee et al., 2018). Providing people injured in freeway crashes with in-time first aid can help reduce severe injury severity.

Although not significant in the SMNL model, wind speed is shown to have significant impacts on freeway injury severity in the MNL model. As shown in Table 4, a one-unit increase in max *wind speed* leads to a 0.7% increase in the probability of *no injury*, a 0.2% probability increase for *minor injury*, and a 0.9% probability decrease in *severe injury and fatality*. Although an increase in wind speed does not increase the probability of severe injury and fatality, the increasing risk of minor injury is still worth noting. As pointed out by previous research, an increase in wind speed might increase the difficulties in handling large vehicles and thus cause a higher possibility for overturning crashes (Hou et al., 2018; Young & Liesman, 2007).

Regarding the vehicle types involved in the accident, when a coach is involved, the possibility for *minor injury* increases 9.54%, and the possibilities for *no injury* and *severe injury and fatality* decrease 8.72% and 0.82%, respectively. When a truck is involved in the accident, the risks for *minor injury* and *severe injury and fatality* increase by 5.8% and 4.3%, respectively. This result is in line with previous studies (Huang et al., 2011; Zeng et al., 2016). A possible reason might be that coaches and trucks have high destructive capabilities, leading to greater jeopardy to other vehicles that are involved in the same accident. In addition, *other vehicles* are shown to have significant impacts on injury severity in the MNL model, indicating that if the involved vehicle includes other vehicles, such as trailers, the possibility of minor injury increases by 5.90%, with a 5.40% decrease in the possibility of no injury and a 0.50% decrease in severe injury and fatality. A possible reason might be the high destructive capabilities of these types of vehicles. Once the accident occurs, the other cars in the same accident are greatly damaged, thus increasing the possibility for minor injury.

Whether the involved vehicle is local has a significant effect on freeway injury severity. When a nonlocal vehicle is involved in the crash, the possibility for *minor injury* increases by 4%, while the possibilities for *no injury* and *severe injury and fatality* decrease by 3.65% and 0.35%, respectively. A plausible reason for the increased risk of minor injury is that nonlocal drivers are not familiar with the local freeway road alignment. A nonlocal driver might have to pay much more attention to finding and confirming the route and entrance position and thus have increased risk of becoming injured.

Regarding crash time, compared to the afternoon, a crash occurring before dawn is accompanied by more severe outcomes. As shown in the marginal effect result of the SMNL model, compared to that happening before dawn, if a crash occurs in the afternoon, the possibility for *no injury* increases by 2.7%, the possibility for *minor injury* increases by 0.9%, and the possibility of *severe injury and fatality* decreases by 3.6%. A plausible reason for the decreasing probability of *severe injury and fatality* is the high risk for fatigued/drowsy driving before dawn and the better visibility in the afternoon. Also, in the afternoon, with good visibility, it is easy for drivers to track hazards. In comparison, before dawn, drivers tend to feel tired and sleepy when driving, and the light is dim at this time period, which leads to the risk enhancement for severe injury. Meanwhile, the increasing probability of *minor injury* for afternoon is still worth noting and the possible reason is that the strong ultraviolet and glare in the afternoon would make drivers get dizzy or distracted.

Additionally, *crash type* has a tight relationship with freeway injury severity. According to the model results of the SMNL model, compared to a multi-vehicle crash, if a crash is a single-vehicle crash, the possibility for no injury is enhanced by 16.11%, while the possibility for minor injury decreases by 12.42%, and the possibility for severe injury and fatality decreases by 3.69%. A possible reason is that a multi-vehicle crash is the interaction of several vehicles at high speed, and people tend to get injured more easily than in single-vehicle crash.

5. Conclusion and future research directions

This paper investigates the relationship between freeway crash injury severity and hazardous factors incorporating real-time weather factors. In addition, with the purpose of capturing the spatial correlation and improving model fitting, this paper proposes a new method, the spatial multinomial logit (SMNL) model, which introduces conditional autoregressive priors to the traditional multinomial logit (MNL) model, to fit the freeway crash injury data.

For the model comparison result, first, the significant result of the spatial terms in the SMNL model indicates that there exists spatial correlation in the freeway crash data. Second, although the DIC value of the SMNL model is higher than that of the traditional MNL model, the lower $\bar{D}$ value and better classification accuracy of the whole dataset suggest that the SMNL model fits the data better. Moreover, the SMNL model has better classification accuracy for *minor injury* and *severe injury and fatality*, indicating that the results of the SMNL model provide more useful information for freeway injury prevention.

Regarding the parameter estimate result of the SMNL model, it could be found that several risk factors have significant effects on the freeway injury severity. Specifically, the increase in vertical grade, visibility enhancement, longer EMS response time, truck or coach involvement, nonlocal vehicle involvement, crash occurrence before dawn, and multi-vehicle crashes are accompanied by a higher risk for crash injury.

According to the model results, practical countermeasures could be put forward to prevent the occurrence of severe crash outcomes. For instance, in regard to freeway roadway design, steep slopes should be avoided if possible, and in-lane rumble strips should be recommended for steep down-slope segments. For traffic management, the traffic volume proportion of large vehicles could be limited when the wind speed exceeds a certain grade. For the improvement of EMS, real-time accident detection technology is recommended to replace the traditional incident detection method, and the EMS facility location should be set optimally with the goal of the shortest rescue arrival time. For the design of vehicles, the construction of trucks or coaches could be improved to reduce the aggressiveness of the vehicles, and tires with good heat resistance are also encouraged to be used by large vehicles. In addition, drivers should be encouraged to use sunglasses when driving in the afternoon, to reduce the adverse effect on driving behavior brought by the strong sunlight.

As for the limitation, the data are only collected from one freeway, and the data sample size is of limited. For future research, the data from more freeways will be collected. Besides, how the interaction effects between weather factors and other factors (such as roadway factors, driver factors, and vehicle factors) affect freeway crash injury severity will be investigated and more high-risk scenes based on the results will be identified in the future. With this research, more targeted countermeasures could be put forward.

Acknowledgements

This study is supported by (1) the International Science & Technology Cooperation Program of China under grant no. 2017YFE0134500, (2) the Natural Science Foundation of China under grant no. 71801095 and (3) the China Scholarship Council (CSC).

References

Carson, J., & Mannering, F. (2001). The effect of ice warning signs on ice-accident frequencies and severities. *Accident Analysis & Prevention, 33*(1), 99–109. https://doi.org/10.1016/S0001-4575(00)00020-8.

Chen, Z., & Fan, W. (David) (2019). A multinomial logit model of pedestrian-vehicle crash severity in North Carolina. *International Journal of Transportation Science and Technology, 8*(1), 43–52. https://doi.org/10.1016/J.IJTST.2018.10.001.

Christoforou, Z., Cohen, S., & Karlaftis, M. G. (2010). Vehicle occupant injury severity on highways: An empirical investigation. *Accident Analysis & Prevention, 42*(6), 1606–1620. https://doi.org/10.1016/J.AAP.2010.03.019.

Golob, T. F., Recker, W. W., & Leonard, J. D. (1987). An analysis of the severity and incident duration of truck-involved freeway accidents. *Accident Analysis & Prevention, 19*(5), 375–395. https://doi.org/10.1016/0001-4575(87)90023-6.

Gonzalez, R. P., Cummings, G. R., Phelan, H. A., Mulekar, M. S., & Rodning, C. B. (2009). Does increased emergency medical services prehospital time affect patient mortality in rural motor vehicle crashes? A statewide analysis. *The American Journal of Surgery, 197*(1), 30–34. https://doi.org/10.1016/J.AMJSURG.2007.11.018.

Haleem, K., & Gan, A. (2013). Effect of driver's age and side of impact on crash severity along urban freeways: A mixed logit approach. *Journal of Safety Research, 46*, 67–76. https://doi.org/10.1016/J.JSR.2013.04.002.

Hou, Q., Huo, X., Leng, J., & Cheng, Y. (2019). Examination of driver injury severity in freeway single-vehicle crashes using a mixed logit model with heterogeneity-in-means. *Physica A: Statistical Mechanics and its Applications, 531*. https://doi.org/10.1016/j.physa:2019.121760 121760.

Hou, Q., Tarko, A. P., & Meng, X. (2018). Investigating factors of crash frequency with random effects and random parameters models: New insights from Chinese freeway study. *Accident Analysis & Prevention, 120*, 1–12. https://doi.org/10.1016/J.AAP.2018.07.010.

Huang, H., Siddiqui, C., & Abdel-Aty, M. (2011). Indexing crash worthiness and crash aggressivity by vehicle type. *Accident Analysis & Prevention, 43*(4), 1364–1370. https://doi.org/10.1016/J.AAP.2011.02.010.

Huang, H., Song, B., Xu, P., Zeng, Q., Lee, J., & Abdel-Aty, M. (2016). Macro and micro models for zonal crash prediction with application in hot zones identification. *Journal of Transport Geography, 54*, 248–256. https://doi.org/10.1016/J.JTRANGEO.2016.06.012.

Jeihani, M., & Banerjee, S. (2018). Drivers' behavior analysis under reduced visibility conditions using a driving simulator. *Journal of Traffic and Logistics Engineering, 6*(2).

Jung, S., Qin, X., & Noyce, D. A. (2010). Rainfall effect on single-vehicle crash severities using polychotomous response models. *Accident Analysis & Prevention, 42*(1), 213–224. https://doi.org/10.1016/J.AAP.2009.07.020.

Lee, J., Abdel-Aty, M., Cai, Q., & Wang, L. (2018). Effects of emergency medical services times on traffic injury severity: A random effects ordered probit approach. *Traffic Injury Prevention, 19*(6), 577–581. https://doi.org/10.1080/15389588.2018.1468889.

Lu, J. J., Xing, Y., Wang, C., & Cai, X. (2016). Risk factors affecting the severity of traffic accidents at Shanghai river-crossing tunnel. *Traffic Injury Prevention, 17*(2), 176–180. https://doi.org/10.1080/15389588.2015.1051222.

Ma, X., Chen, S., & Chen, F. (2017). Multivariate space-time modeling of crash frequencies by injury severity levels. *Analytic Methods in Accident Research, 15*, 29–40. https://doi.org/10.1016/J.AMAR.2017.06.001.

Ma, Z., Chien, S.-J., Dong, C., Hu, D., & Xu, T. (2016). Exploring factors affecting injury severity of crashes in freeway tunnels. *Tunnelling and Underground Space Technology, 59*, 100–104. https://doi.org/10.1016/J.TUST.2016.06.013.

Malyshkina, N. V., & Mannering, F. L. (2010). Empirical assessment of the impact of highway design exceptions on the frequency and severity of vehicle accidents. *Accident Analysis & Prevention, 42*(1), 131–139. https://doi.org/10.1016/J.AAP.2009.07.013.

Meng, F., Xu, P., Wong, S. C., Huang, H., & Li, Y. C. (2017). Occupant-level injury severity analyses for taxis in Hong Kong: A Bayesian space-time logistic model. *Accident Analysis & Prevention, 108*, 297–307. https://doi.org/10.1016/J.AAP.2017.08.010.

Mergia, W. Y., Eustace, D., Chimba, D., & Qumsiyeh, M. (2013). Exploring factors contributing to injury severity at freeway merging and diverging locations in Ohio. *Accident Analysis & Prevention, 55*, 202–210. https://doi.org/10.1016/J.AAP.2013.03.008.

Mueller, A. S., & Trick, L. M. (2012). Driving in fog: The effects of driving experience and visibility on speed compensation and hazard avoidance. *Accident Analysis & Prevention, 48*, 472–479. https://doi.org/10.1016/J.AAP.2012.03.003.

Naik, B., Tung, L.-W., Zhao, S., & Khattak, A. J. (2016). Weather impacts on single-vehicle truck crash injury severity. *Journal of Safety Research, 58*, 57–65. https://doi.org/10.1016/J.JSR.2016.06.005.

Savolainen, P., & Mannering, F. (2007). Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes. *Accident Analysis & Prevention, 39*(5), 955–963. https://doi.org/10.1016/J.AAP.2006.12.016.

Scott Long, J. (1997). Regression models for categorical and limited dependent variables. *Advanced Quantitative Techniques in the Social Sciences, 7*.

Shankar, V., & Mannering, F. (1996). An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity. *Journal of Safety Research, 27*(3), 183–194. https://doi.org/10.1016/0022-4375(96)00010-2.

Shankar, V., Mannering, F., & Barfield, W. (1995). Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. *Accident Analysis & Prevention, 27*(3), 371–389. https://doi.org/10.1016/0001-4575(94)00078-Z.

Shankar, V., Mannering, F., & Barfield, W. (1996). Statistical analysis of accident severity on rural freeways. *Accident Analysis & Prevention, 28*(3), 391–401. https://doi.org/10.1016/0001-4575(96)00009-7.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *J Royal Statistical Soc B, 64*(4), 583–639. https://doi.org/10.1111/1467-9868.00353.

Theofilatos, A., & Yannis, G. (2014). A review of the effect of traffic and weather characteristics on road safety. *Accident Analysis & Prevention, 72*, 244–256. https://doi.org/10.1016/j.aap.2014.06.017.

Theofilatos, A., & Yannis, G. (2017). Investigation of powered 2-wheeler accident involvement in urban arterials by considering real-time traffic and weather data. *Traffic Injury Prevention, 18*(3), 293–298. https://doi.org/10.1080/15389588.2016.1198871.

Ulfarsson, G. F., & Mannering, F. L. (2004). Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. *Accident Analysis & Prevention, 36*(2), 135–147. https://doi.org/10.1016/S0001-4575(02)00135-5.

Xu, X., Xie, S., Wong, S. C., Xu, P., Huang, H., & Pei, X. (2016). Severity of pedestrian injuries due to traffic crashes at signalized intersections in Hong Kong: A Bayesian spatial logit model. *Journal of Advanced Transportation, 50*(8), 2015–2028. https://doi.org/10.1002/atr.1442.

Ye, F., & Lord, D. (2014). Comparing three commonly used crash severity models on sample size requirements: Multinomial logit, ordered probit and mixed logit models. *Analytic Methods in Accident Research, 1*, 72–85. https://doi.org/10.1016/J.AMAR.2013.03.001.

Young, R. K., & Liesman, J. (2007). Estimating the relationship between measured wind speed and overturning truck crashes using a binary logit model. *Accident Analysis & Prevention, 39*(3), 574–580. https://doi.org/10.1016/J.AAP.2006.10.002.

Yu, R., & Abdel-Aty, M. (2014). Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. *Safety Science, 63*, 50–56. https://doi.org/10.1016/J.SSCI.2013.10.012.

Zeng, Q., Gu, W., Zhang, X., Wen, H., Lee, J., & Hao, W. (2019). Analyzing freeway crash severity using a Bayesian spatial generalized ordered logit model with conditional autoregressive priors. *Accident Analysis & Prevention, 127*, 87–95. https://doi.org/10.1016/J.AAP.2019.02.029.

Zeng, Q., Hao, W., & Chen, F. (2020). Investigating the impacts of real-time weather conditions on freeway crash severity: A Bayesian spatial analysis. *International Journal of Environmental Research and Public Health, 17*, 2768. https://doi.org/10.3390/ijerph17082768.

Zeng, Q., & Huang, H. (2014a). Bayesian spatial joint modeling of traffic crashes on an urban road network. *Accident Analysis & Prevention, 67*, 105–112. https://doi.org/10.1016/J.AAP.2014.02.018.

Zeng, Q., & Huang, H. (2014b). A stable and optimized neural network model for crash injury severity prediction. *Accident Analysis & Prevention, 73*, 351–358. https://doi.org/10.1016/J.AAP.2014.09.006.

Zeng, Q., Wen, H., & Huang, H. (2016). The interactive effect on injury severity of driver-vehicle units in two-vehicle crashes. *Journal of Safety Research, 59*, 105–111. https://doi.org/10.1016/J.JSR.2016.10.005.

Zeng, Q., Wen, H., Huang, H., & Abdel-Aty, M. (2017). A Bayesian spatial random parameters Tobit model for analyzing crash rates on roadway segments. *Accident Analysis & Prevention, 100*, 37–43. https://doi.org/10.1016/J.AAP.2016.12.023.

Zeng, Q., Wen, H., Wong, S. C., Huang, H., Guo, Q., & Pei, X. (2020). Spatial joint analysis for zonal daytime and nighttime crash frequencies using a Bayesian bivariate conditional autoregressive model. *Journal of Transportation Safety & Security, 12*(4), 566–585. https://doi.org/10.1080/19439962.2018.1516259.

Zhai, X., Huang, H., Sze, N. N., Song, Z., & Hon, K. K. (2019). Diagnostic analysis of the effects of weather condition on pedestrian crash severity. *Accident Analysis & Prevention, 122*, 318–324. https://doi.org/10.1016/J.AAP.2018.10.017.

Zhang, Y., Li, Z., Liu, P., & Zha, L. (2011). Exploring contributing factors to crash injury severity at freeway diverge areas using ordered probit model. *Procedia Engineering, 21*, 178–185.

Xuan Zhang is a Ph.D. candidate in transportation planning and management at South China University of Technology, China. She received her B.E. degree in transportation from South China University of Technology, China. Her research interest is mainly on traffic safety.

Huiying Wen received her Ph.D. degree in transportation engineering from South China University of Technology, China, in 2007. She is a professor, a deputy dean of School of Civil Engineering and Transportation, South China University of Technology. Her research interests refer to traffic safety analysis, transportation planning and management.

Toshiyuki Yamamoto, Ph.D., is a Professor in Institute of Materials and Systems for Sustainability, Nagoya University, Japan. His primary research areas of interest include traffic safety, ITS technologies for environmentally sustainable transport, household vehicle ownership, and activity analysis.

Qiang Zeng received his B.E. and Ph.D. degrees in transportation engineering from Central South University, China, in 2010 and 2015 respectively. Now, he is an associate professor in the School of Civil Engineering and Transportation, South China University of Technology, China. His research interests cover traffic safety analysis, transportation system modeling and public transit operation & management.

Journal of Safety Research 76 (2021) 301–313

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Investigating the impacts of crash prediction models on quantifying safety effectiveness of Adaptive Signal Control Systems

Weimin Jin [a,*], Mashrur Chowdhury [a], Sakib Mahmud Khan [b], Patrick Gerard [c]

[a] Glenn Department of Civil Engineering, Clemson University, Clemson, SC 29634, USA
[b] Center for Connected Multimodal Mobility, Clemson University, Clemson, SC 29634, USA
[c] School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC 29634, USA

ARTICLE INFO

Article history:
Received 2 March 2020
Received in revised form 25 June 2020
Accepted 12 November 2020
Available online 16 December 2020

Keywords:
Safety effectiveness
Empirical Bayes
Fully Bayesian
Spatial model
Adaptive signal control

ABSTRACT

Introduction: Adaptive Signal Control System (ASCS) can improve both operational and safety benefits at signalized corridors. *Methods:* This paper develops a series of models accounting for model forms and possible predictors and implements these models in Empirical Bayes (EB) and Fully Bayesian (FB) frameworks for ASCS safety evaluation studies. Different models are validated in terms of the ability to reduce the potential bias and variance of prediction and improve the safety effectiveness estimation accuracy using real-world crash data from non-ASCS sites. This paper then develops the safety effectiveness of ASCS at six different corridors with a total of 65 signalized intersections with the same type of ASCS, in South Carolina. *Results:* Validation results show that the FB model that accounts for traffic volume, roadway geometric features, year factor, and spatial effects shows the best performance among all models. The study findings reveal that ASCS reduces crash frequencies in the total crash, fatal and injury crash, and angle crash for most of the intersections. The safety effectiveness of ASCS varies with different intersection features (i.e., AADT at major streets, number of legs at an intersection, the number of through lanes on major streets, the number of access points on minor streets, and the speed limit at major streets). *Conclusions:* ASCS is associated with crash reductions, and its safety effects vary with different intersection features. *Practical Applications:* The findings of this research encourage more ASCS deployments and provide insights into selecting ASCS deployment sites for reducing crashes considering the variation of the safety effectiveness of ASCS.

1. Introduction

Safety improvements at intersections have become one of 22 key domains in the American Association of State Highway and Transportation Officials Strategic Highway Safety Plan (Antonucci et al., 2004). The goal of this plan is to achieve a decrease in the frequency and severity of crashes at signalized intersections. Transportation agencies have been advancing new approaches and technologies to improve safety benefits at signalized intersections.

Adaptive Signal Control System (ASCS) is typically deployed at intersections to improve operational performance, such as travel time and traffic delay. The ASCS requires detectors such as loop detectors, and a communication network that allows for communicating with the local traffic controllers and/or the server. Compared to the conventional time of day signal control systems (i.e., pre-timed signal control and actuated signal control) with predefined signal plans (usually re-adjusted every two years), ASCS can change the signal timings (i.e., phase splits, phase sequence, offsets, and cycle length) in real-time to accommodate fluctuating traffic demand at intersections. Also, ASCS can adjust offsets to coordinate several intersections along a corridor, thus lead to fewer traffic stops. Significant operational benefits of ASCS in both corridors and intersections have been documented (Eghtedari, 2005; Elkins et al., 2012; Fontaine et al., 2015; Kergaye et al., 2009; Khattak, 2016; Khattak et al., 2020; So et al., 2014). ASCS can potentially improve the traffic operations, which in turn will improve the safety of signalized intersections and corridors.

Past studies have implemented the Empirical Bayes (EB) framework with the Poisson-Gamma model into ASCS safety evaluation studies (Jesus & Benekohal, 2019; Khattak, 2016). However, previous studies have not applied the Fully Bayesian (FB) framework with the Poisson-Lognormal model for any ASCS safety evaluation. More specifically, spatial correlations could exist in neighboring intersections along a corridor with ASCS. However, no studies have

* Corresponding author.
E-mail addresses: weiminj@g.clemson.edu (W. Jin), pgerard@clemson.edu (P. Gerard).

https://doi.org/10.1016/j.jsr.2020.11.003
0022-4375/

implemented spatial models in the safety evaluation of ASCS. Moreover, previous studies have not evaluated the performance of different crash prediction models in quantifying the safety effectiveness of ASCS in the EB and FB before-and-after studies.

To fill the above research gaps in ASCS safety evaluation, we: (a) implement the Poisson-Lognormal model and the spatial model into the ASCS safety evaluation; and (b) investigate how different crash prediction models impact the estimator of the safety effectiveness of ASCS in the EB and FB before-and-after studies. A series of models, including the Poisson-Lognormal models, Poisson-Gamma models, and spatial models, are compared and evaluated. Traffic volume, roadway geometric features, year factor, and spatial effect are used to produce different sets of models. The intersections in this study have the same ASCS type deployed. The study focuses on evaluating the safety effectiveness of the particular ASCS (which this paper refers to as "ASCS") without considering the variations between multiple ASCS types. ASCS effects may vary across sites due to specific features of the sites that are deployed with ASCS. To explore the variations in ASCS effect across sites, the study evaluates the safety effectiveness of ASCS for each corridor and each intersection.

2. Literature review

The following subsections discuss the crash prediction model and safety evaluation related studies of ASCS.

2.1. Crash prediction model

This subsection reviews the characteristics of the crash prediction models. In the FB method, the Bayesian models used to estimate the safety performance are similar to the concept of the Safety Performance Function (SPF) used in the EB method. This paper uses the same term "crash prediction model" for the convenience of discussion, instead of the Bayesian models in the FB methods and the SPF in the EB method.

2.1.1. Poisson-Gamma and Poisson-Lognormal Model

In general, there are two main types of models used in the estimation of crash frequency: (a) Poisson-Gamma, and (b) Poisson-Lognormal.

- *Poisson-Gamma Model*

When the Poisson mean is assumed to follow a gamma distribution, the Poisson-Gamma mix distribution results in Negative Binomial (NB) distribution (Carriquiry & Pawlovich, 2004; Khazraee et al., 2018), with Maximum Likelihood Estimation (MLE) used for parameter estimation. NB models have been widely used by many researchers (Elvik et al., 2017; Hauer, 1997; Hauer et al., 2002; Hovey & Chowdhury, 2005; Høye, 2015). In the EB framework, the NB model is used to account for the over-dispersion (i.e., the variance is much larger than the mean) of crash data.

- *Poisson-Lognormal Model*

When the Poisson mean is assumed to have a lognormal distribution, the Poisson-Lognormal model results in an unclosed form of the marginal distribution, which is difficult to handle using the MLE method. The Poisson-Lognormal model is typically integrated into the FB framework. The posterior distribution of the parameters of the Poisson-Lognormal model can be obtained using Markov Chain Monte Carlo (MCMC) simulation (Khazraee et al., 2018).

2.1.2. Spatial models

Spatial effects can be introduced into a Poisson-Lognormal model to consider the spatial correlation of adjacent road entities (Cai et al., 2018). Although many studies (Barua et al., 2016; Jonathan et al., 2016) have accounted for spatial effects in the development of crash prediction models, few studies (Sacchi et al., 2016) implement the spatial model in a before-and-after safety study. Spatial models can be integrated into the FB method but cannot be in current EB methods for safety evaluation (Gross et al., 2010). The assumption of the non-spatial models (i.e., Poisson-Gamma model and Poisson-Lognormal model) is that crashes are independent across sites. This assumption will be violated if the spatial correlation between sites within neighborhoods exists.

On the other hand, neighboring sites may share similar traffic and road conditions, driver behavior, and weather condition. As a result, it may result in similar safety levels for neighboring sites. Spatial effects usually exist, for example, among the adjacent intersections (which is the case of this study), adjacent corridors (Li & Wang, 2017), and the adjacent zone sharing the same border (Cai et al., 2018).

2.2. Safety evaluation of Adaptive Signal Control Systems

Safety benefits of ASCS have been demonstrated in recent studies. Fontaine et al. (2015) have evaluated the safety effects of InSync, a type of ASCS, for different corridors in Virginia using an EB before-and-after study. Based on the analysis, the authors have found that crashes are reduced by 17% due to ASCS. Dutta et al. (2010) have studied crash data for one type of ASCS (i.e., SCATS) and fixed-time signal control systems for two corridors in Michigan. The authors (Dutta et al., 2010) have evaluated the change in the crash rate before and after the ASCS deployment. The authors have found that the total crash rate is reduced by 6% after installing ASCS. The incapacitating injury crashes are reduced by 22% after ASCS deployment. The most significant improvement is found for non-incapacitating injury crashes, which is reduced by 35%. Fink et al. (2016) have studied the safety impacts of SCATS installed at signalized intersections in Oakland County. The authors have performed a cross-sectional study using data from 498 signalized intersections and found a reduction of 19.3% in angle crashes associated with SCATS. This study found that SCATS does not significantly reduce incapacitating injuries or fatality (Fink et al., 2016). Khattak (2016) evaluated 41 intersections in Pennsylvania where SURTRAC and InSync are installed. The author has implemented an EB before-and-after safety study and computed Crash Modification Factors (CMF) for total crashes, and fatal and injury crashes. The author found reductions of 34% and 45% in total crashes and fatal and injury crashes, respectively, due to ASCS. Khattak et al. (2019) have examined the impact of ASCS on crash severity. The authors have found that one type of ASCS (the name of the ASCS type is not mentioned in the paper) decreases the probability of minor injury and severe plus moderate crashes by 10.36% and 11.70%, respectively, while another type of ASCS (the name of the ASCS type is not mentioned in the paper) decreases the probability of minor injury and severe plus moderate crashes by 6.92% and 4.39%, respectively. Jin et al. (2021) have investigated the effects of ASCS on crash severity. The authors (Jin et al., 2021) have found that ASCS is associated with lower crash severity. The effects of ASCS on crash severity vary with different intersection and corridor features.

ASCS is not always found to reduce crashes in a statistically significant manner. Jesus and Benekohal (2019) have implemented the EB method to determine the safety effectiveness of the ASCS. The authors (Jesus & Benekohal, 2019) have found that the CMF of ASCS for fatal and injury crashes is 0.67 (CMF less than 1 indicates that ASCS reduces crashes), which is not statistically significant at a 0.05 significance level. CMFs of property damage only and total crashes are close to one, which indicates no crash reduction due to ASCS. The CMF for fatal, incapacitating injury, and non-incapacitating injury combined is 0.68, which is not significant at a 0.05 significance level. The angle, rear-end, incapacitating injury, and reported/not evident injury (this includes momentary unconsciousness, claims of no evident injuries, limping, complaints of pain, nausea, hysteria) crashes show insignificant reductions.

3. Method

This section first discusses model forms in the development of crash prediction models in the EB and FB before-and-after study procedures. Then, this section provides a validation procedure that uses two criteria to validate possible models: (a) the potential bias and variance of prediction, and (b) the estimation accuracy of safety effectiveness.

3.1. Model development and evaluation procedure

This subsection introduces the models that would be incorporated into the EB and FB before-and-after study procedures. Traffic volume, roadway geometric features (e.g., the number of access points at an intersection, and the number of exclusive left-turn lanes, right-turn lanes, and through lanes on major or minor streets), year factor, and spatial effects are used to produce different sets of the models. For each model, four crash types of interest are accounted for: total crash, fatal and injury (F + I) crash, rear-end crash, and angle crash. Two primary forms of models, Poisson-Gamma and Poisson-Lognormal, are introduced. A spatial model is also used with a Poisson-Lognormal model in this study to account for the spatial effects existing in the investigated sites. Model 1, Model 2, and Model 3 are implemented within the EB framework. Model 4A, Model 4B, Model 5A, Model 5B, Model 6A, and Model 6B are implemented within the FB framework.

3.1.1. EB Models

3.1.1.1. EB Model development. A general Poisson-Gamma model with two tiers is expressed as the following:

$$y_{m,it} \sim Poisson(\lambda_{m,it}) \tag{1}$$

$$\lambda_{m,it} \sim Gamma(\alpha, \phi) \tag{2}$$

where, $y_{m,it}$ is the observed crash frequency at an intersection i ($i = 1, 2, \ldots, 65$ for a total of 65 intersections considered in this study) on the corridor $m (m = 1, 2, \ldots, 6$ for a total of six corridors considered in this study) in a given year t ($t = 2011, 2012, \ldots,$ 2018); $\lambda_{m,it}$ is the Poisson mean. The expectation of $\lambda_{m,it}$, $E(\lambda_{m,it})$ is the expected yearly number of crashes at an intersection i on the corridor m in the year t for a specified crash type (i.e., total crash, F + I, rear-end, or angle crash). α is the shape parameter of Gamma distribution, and ϕ is the inverse scale parameter (i.e., rate parameter) of the Gamma distribution.

Three crash prediction models (called SPFs in the EB framework) are specified in terms of different explanatory variables. Model 1 and Model 2 account for the year factor by introducing annual multipliers. The year factor is often introduced into the crash prediction model to account for temporal variation of crash expectation, which accounts for possible unobserved factors such as weather conditions, road conditions, and vehicle technology improvements (Persaud et al., 2010). Model 3 accounts for the year factor by introducing the year variable as one of the explanatory variables in the model. Model 1 includes an annual multiplier, and Annual Average Daily Traffic (AADT) without considering the difference in roadway geometric features. Model 2 includes an annual multiplier, AADT, and roadway geometric features. Model 3 includes AADT, roadway geometric features, and the year factor.

Model 1 (AADT + Annual multipliers):

$$E(\lambda_{m,it}) = a_{m,t}\exp(\beta_{m,0} + \beta_{m,maj-aadt}log(majorAADT_{m,it}) + \beta_{m,\min-aadt}log(minorAADT_{m,it})) \tag{3}$$

Model 2 (AADT + Roadway factor + Annual multipliers):

$$E(\lambda_{m,it}) = a_{m,t}\exp(\beta_{m,0} + \beta_{m,maj-aadt}log(majorAADT_{m,it}) + \beta_{m,\min-aadt}log(minorAADT_{m,it}) + \sum_{n=1,j=1}^{Q} \beta_{m,j}X_{mn,it}) \tag{4}$$

Model 3 (AADT + Roadway factor + Year):

$$E(\lambda_{m,it}) = \exp(\beta_{m,0} + \beta_{m,maj-aadt}log(majorAADT_{m,it}) + \beta_{m,\min-aadt}log(minorAADT_{m,it}) + \sum_{n=1,j=1}^{Q} \beta_{m,j}X_{mn,it} + \beta_{m,T}T_{m,it}) \tag{5}$$

where, $majorAADT_{m,it}$ is AADT of major roads at the intersection i on the corridor m in a given year t; $minorAADT_{m,it}$ is AADT of minor streets at the intersection i on the corridor m in a given year t; $X_{mn,it}$ is the n^{th} explanatory variable of roadway geometric features (e.g., the number of exclusive left-turn, right-turn lane(s) and through lane(s) on major or minor streets and the number of access point(s) at an intersection) for the intersection i in a given year t; Q is the total number of explanatory variables of roadway geometric features; $T_{m,it}$ is the year factor which is numeric, for example, 0 if year is 2011, 1 if year is 2012, and so on; $\beta_{m,T}$ is the coefficient for the year factor of Model 3; $\beta_{m,maj-aadt}$is the coefficient for AADT of major roads; $\beta_{m,\min-aadt}$is the coefficient for AADT of minor streets; $\beta_{m,0}$ is the intercept and $\beta_{m,j}$ is the j^{th} coefficient for roadway geometric features in Model 2 and Model 3. $a_{m,t}$ is the annual multiplier which is obtained by dividing the sum of predicted number of crashes in a given year t by the sum of observed crashes in a given year t after the EB models are fitted.

3.1.1.2. EB Model estimation and evaluation. EB model estimation is performed in the R software by calling the R package "MASS." Potential multicollinearity MC issues are checked by evaluating the Variance Inflation Factor (VIF) statistic. VIF values greater than 10 are used to check whether MC is of concern (O'Brien 2007). Using this criterion, the authors find that no MC issues exist among the explanatory variables used in this study. Akaike's Information Criterion (AIC) is used to select the set of variables used in the regression models (Bumham & Anderson, 2002). The best-fitted model is found with the lowest AIC. For example, roadway geometric features have some variables, including the number of exclusive left-turn lanes, right-turn lanes, and through lanes on major or minor streets and the number of the access points at an intersection. After model selection based on AIC, only a few roadway geometric variables will be kept.

3.1.1.3. EB before-and-after evaluation procedure. The expected number of crashes in the before period E_b, is obtained by combining two different information sources: (1) the observed crash data for a site, O_b, and (2) the sum of the predicted number of crashes during the before period, P_b, estimated by the crash prediction models (i.e., Model 1, Model 2, and Model 3) for the individual site. E_b is obtained by using the following equation (Hauer, 1997; Persaud & Lyon, 2007),

$$E_b = wP_b + (1 - w)O_b \tag{6}$$

The weight factor is estimated from P_b and ψ, which are estimated from the SPF development,

$$w = \frac{1}{1 + P_b/\psi} \tag{7}$$

where ψ is the value of the dispersion parameter obtained by the NB regression-based SPF.

A correction factor that accounts for the length of the after period, changes in traffic volumes, and changes in roadway geometric characteristics is multiplied with $E_b$to obtain theE_a. This factor is the ratio of the sum of the after-period SPF predictions, P_a and the sum of the before-period SPF predictions, P_b. Thus, E_a can be obtained below,

$$E_a = E_b \frac{P_a}{P_b} \tag{8}$$

The observed number of crashes at a site with treatment during the after period (O_a) is then compared to the expected number of crashes on the same site (E_a), which is the expected number of crashes that would have occurred if the treatment had not been implemented. An estimate of the index of safety effectiveness of treatment, θ, is:

$$\theta = \frac{\sum_{all} O_a / \sum_{all} E_a}{1 + Var(\sum_{all} E_a)/(\sum_{all} E_a)^2} \tag{9}$$

$$Var\left(\sum_{all} E_a\right) = \sum_{all}\left[(P_a/P_b)^2 E_b (1-w)\right] \tag{10}$$

where, $\sum_{all} O_a$ is the summation of O_a for all studied sites; $\sum_{all} E_a$ is the summation of E_a for all studied sites.

The estimated percentage of reduction in crashes is $100(1-\theta)$. For example, a value of $\theta = 0.45$ indicates a 55% decrease in crashes with treatment. The uncertainty of the index of effectiveness (i.e., standard deviation) is calculated by taking the square root of the variance of θ. The variance of θ is (Hauer, 1997; Persaud & Lyon, 2007):

$$Var(\theta) = \frac{\theta^2\left(\frac{Var(\sum_{all} O_a)}{(\sum_{all} O_a)^2} + \frac{Var(\sum_{all} E_a)}{(\sum_{all} E_a)^2}\right)}{\left(1 + \frac{Var(\sum_{all} E_a)}{(\sum_{all} E_a)^2}\right)^2} \tag{11}$$

In the Eq. (10), the assumption is that the ratio P_a to P_b is a constant variable, not a random variable, which would affect the Eq. (9) and Eq. (11) containing the term $Var(\sum_{all} E_a)$.

3.1.2. FB models

3.1.2.1. FB model development.

A general Poisson-Lognormal model is introduced with multiple hierarchical levels in the following:

$$y_{m,it} \sim Poisson(\lambda_{m,it}) \tag{12}$$

$$\log(\lambda_{m,it}) = \sum_{j=0}^{p} \beta_{mj,B} B_{mj,it} + \varepsilon_{m,it} \tag{13}$$

$$\varepsilon_{m,it} \sim Normal(0, \sigma_\varepsilon^2) \tag{14}$$

$$\beta_{mj,B} \sim Normal(0, \sigma_{\beta,j}^2) \tag{15}$$

where, $y_{m,it}$ is the observed crash frequency at the intersection i on the corridor m in a given year t; $\lambda_{m,it}$ is the Poisson mean. $B_{mj,it}$ is the explanatory variable in the model. $\beta_{mj,B}$ is the j^{th} coefficient for the explanatory variable in the model. P is the total number of explanatory variables. The distribution of parameters such as $\lambda_{m,it}$,$\beta_{mj,B}$, and $\varepsilon_{m,it}$ in the model is evaluated based on the estimation of the posterior distribution of these parameters using the FB approach. In the FB models, $\lambda_{m,it}$ is the site-specific expected crash frequency, and each $\lambda_{m,it}$ represents a model parameter. $\varepsilon_{m,it}$ is introduced to account for the variation across intersections and years. σ_ε^2 is assumed to follow a prior Inverse-Gamma (0.001, 0.001) distribution for all models based on previous studies (Cai et al., 2018; Carriquiry & Pawlovich, 2004; Sacchi & Sayed, 2014). $\sigma_{\beta,j}^2$ is set to 1000 for all the prior distributions of $\beta_{mj,B}$ for all models resulting in a non-informative prior distribution for $\beta_{mj,B}$ (Persaud et al., 2010). Consequently, estimation of the posterior distribution of $\beta_{mj,B}$ largely depends on observed data.

Three FB non-spatial models are defined in terms of different explanatory variables. Model 4A and Model 5A introduce a random effect to account for variation caused by the various intersections and years, while Model 6A directly treats the year factor as a covariate in the model. Based on the inclusion of the spatial effect into the models, three different FB spatial models-Model 4B, Model 5B, and Model 6B are developed. A corridor-specific ASCS indicator variable $I_{m,it}$ that labels the after period during which ASCS is installed on the corridor m is included as shown below (1 is the after period; 0 otherwise). $\beta_{m,\mathrm{I}}$ is the coefficient of the ASCS presence indicator variable of the following models. The authors initially included the interaction variables into the model to account for the possible interaction between ASCS and AADT and the interaction between ASCS and roadway geometric features in the model. But the interaction variables are not significant. Thus, the interaction variables are not used for the following models.

Model 4A (AADT):

$$\lambda_{m,it} = \exp(\beta_{m,0} + \beta_{m,maj-aadt} log(majorAADT_{m,it}) + \beta_{m,\min-aadt} log(minorAADT_{m,it}) + \beta_{m,I} I_{m,it} + \varepsilon_{m,it}) \tag{16}$$

Model 4B (AADT + Spatial effect):

$$\lambda_{m,it} = \exp(\beta_{m,0} + \beta_{m,maj-aadt} log(majorAADT_{m,it}) + \beta_{m,\min-aadt} log(minorAADT_{m,it}) + \beta_{m,I} I_{m,it} + \varepsilon_{m,it} + s_{m,i}) \tag{17}$$

Model 5A (AADT + Roadway factor):

$$\lambda_{m,it} = \exp(\beta_{m,0} + \beta_{m,maj-aadt} log(majorAADT_{m,it}) + \beta_{m,\min-aadt} log(minorAADT_{m,it}) + \beta_{m,I} I_{m,it} + \sum_{n=1,j=1}^{Q} \beta_{m,j} X_{mn,it} + \varepsilon_{m,it}) \tag{18}$$

Model 5B (AADT + Roadway factor + Spatial effect):

$$\lambda_{m,it} = \exp(\beta_{m,0} + \beta_{m,maj-aadt} log(majorAADT_{m,it}) + \beta_{m,\min-aadt} log(minorAADT_{m,it}) + \beta_{m,I} I_{m,it} + \sum_{n=1,j=1}^{Q} \beta_{m,j} X_{mn,it} + \varepsilon_{m,it} + s_{m,i}) \tag{19}$$

Model 6A (AADT + Roadway factor + Year):

$$\lambda_{m,it} = \exp(\beta_{m,0} + \beta_{m,maj-aadt} log(majorAADT_{m,it}) + \beta_{m,\min-aadt} log(minorAADT_{m,it}) + \beta_{m,I} I_{m,it} + \sum_{n=1,j=1}^{Q} \beta_{m,j} X_{mn,it} + \beta_{m,T} T_{m,it} + \varepsilon_{m,it}) \tag{20}$$

Model 6B (AADT + Roadway factor + Year + Spatial effect):

$$\lambda_{m,it} = \exp(\beta_{m,0} + \beta_{m,maj-aadt} log(majorAADT_{m,it}) + \beta_{m,\min-aadt} log(minorAADT_{m,it}) + \beta_{m,I} I_{m,it} + \sum_{n=1,j=1}^{Q} \beta_{m,j} X_{mn,it} + \beta_{m,T} T_{m,it} + \varepsilon_{m,it} + s_{m,i}) \tag{21}$$

where, $s_{m,i}$ could be considered as a latent variable that captures the effect of unknown or unmeasured covariates that are assumed spatially structured. The intrinsic Conditional Autoregressive (CAR) model (Besag et al., 1991) is used for estimating $s_{m,i}$, which is given by:

$$s_{m,i}|s_{m,j} \sim Normal\left(\frac{\sum_{j\in\partial_i} w_{ij} s_{m,j}}{\sum_{j\in\partial_i} w_{ij}}, \frac{1}{\tau_s \sum_{j\in\partial_i} w_{ij}}\right),\ j\neq i \tag{22}$$

where ∂_i is the set of intersections adjacent to i;w_{ij} is a spatial proximity weight; τ_s is the precision parameter which is the inverse of the variance. τ_s is assumed to follow a prior Gamma (0.001, 0.001) (Cai et al., 2018). w_{ij} is equal to 1 for $i \in \partial_i$; otherwise, w_{ij} is equal to 0.

3.1.2.2. FB model estimation and evaluation. "OpenBUGS" is open-source software that performs Bayesian inference using the Gibbs sampling algorithm. Bayesian model estimation and MCMC simulation are performed in the R software by calling the R package "R2OpenBUGS." For each FB model, two Markov chains are used in MCMC simulations. Each chain has 200,000 iterations and a total of 20,000 iterations are discarded during the burn-in (i.e., warm-up) period. Bayesian estimation provides posterior probability distributions and Bayesian Credible Intervals (BCI) for statistical inference. Before implementing the estimation of the posterior distribution of parameters of interest, convergence must be checked in the MCMC simulation. As a rule of thumb, Rhat statistics (i.e., scale reduction factor) less than 1.2 (Brooks et al., 1998) is used to identify convergence. Also, viewing graphical summaries and the number of effective samplings (i.e., the number of independent samples drawn from the posterior distribution in the MCMC simulation) for the parameters of interest could help to check the convergence. Deviance Information Criterion (DIC) can be used to determine the best set of predictors for each FB model (Spiegelhalter et al., 2002). In general, differences of more than 10 (DIC value) may suggest that the FB model with lower DIC is preferred (Spiegelhalter et al., 2002). Also, the significance of the spatial effect is evaluated to determine if the spatial effect exits in the crash data.

3.1.2.3. FB before-and-after evaluation procedure. In the FB before-and-after study procedure, Crash Reduction Rate (CRR) is calculated (Lan et al., 2009; Persaud et al., 2010; Yanmaz-Tuzel & Ozbay, 2010), as

$$CRR = 1 - \frac{\sum_{all} O_a}{\sum_{all} \mu_a} \tag{23}$$

$\frac{\sum_{all} O_a}{\sum_{all} \mu_a}$ is similar to the index of the safety effectiveness used in the EB method.

The observed number of crashes at a site with treatment during the after period (O_a) is compared with the expected number of crashes on the same site (μ_a), which is the number of crashes that would have occurred if the treatment had not been implemented. μ_a can be obtained through developing crash prediction models (i.e., Model 4A, Model 4B, Model 5A, Model 5B, Model 6A, and Model 6B) in the FB procedure. $\sum_{all} \mu_a$ is the summation of μ_a for all studied intersections on a corridor across studied years for corridor-specific safety effectiveness calculation or the summation of for a specific intersection across studied years for intersection-specific safety effectiveness calculation.

CRR is obtained directly by MCMC simulation. The uncertainty of CRR can be evaluated with a 95% BCI by MCMC simulation. The significance of CRR can be determined if the 95% BCI does not contain zero.

3.2. Validation of the before-and-after evaluation methods

This section provides a validation procedure that uses two criteria to validate EB and FB models: (a) the potential bias and variance of prediction, and (b) the estimation accuracy of safety effectiveness. In this way, EB and FB models are compared using the same criteria adopted in this study.

3.2.1. Evaluation of potential bias and variance of prediction

Root Mean Square Error (RMSE) is used to compare the potential bias and variance of prediction among different models. RMSE is also used to measure the quality of an estimator and represent the model prediction error and the model goodness of fit. A lower value of RMSE indicates a smaller difference between the estimated value and the actual observed crash frequency for non-ASCS intersections. The equation is shown below:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \sum_{t=1}^{T} (E_{it} - O_{it})^2}{NT}} \tag{24}$$

where, E_{it} is the expected number of the crashes of non-ASCS intersections in an intersection i in the year t; O_{it} is the observed crashes of non-ASCS intersections in an intersection$i$ in the year t; N is the total number of non-ASCS intersections used for validation; T is the total number of years.

In the EB procedure, the expected number of crashes in the subsequent years for a specific intersection can be estimated by multiplying a correction factor due to the difference between the subsequent years and the predecessor year by the expected number of crashes in the predecessor years. For example, the estimated crash in 2012 for an intersection can be obtained by multiplying the correction factor due to the difference between 2011 and 2012 by the expected number of crashes in 2011. Likewise, the estimated crash frequency in 2013, 2014, 2015, 2016, and 2017 can be predicted in this way. In the FB procedure, the expected number of crashes for a specific intersection in a given year can be estimated directly by the MCMC simulation.

3.2.2. Estimation of safety effectiveness of non-ASCS intersections

To evaluate the performance of the candidate models in estimating the safety effectiveness of ASCS, the authors compute and compare the safety effectiveness of ASCS for non-ASCS intersections among different models since no ASCS effect exists for the non-ASCS intersections. So crash reduction percentage for the non-ASCS intersections (i.e., zero) can be deemed as the ground truth. In the EB procedure, the null hypothesis is that the crash reduction percentage is equal to zero, and the alternative hypothesis is that the crash reduction percentage is not equal to zero. In the FB procedure, the significance of the crash reduction percentage is determined if the 95% BCI does not contain zero. To calculate the crash reduction percentage for the non-ASCS intersections, the authors assume that 2011–2014 is the "before period;" 2015–2017 is the "after period" just for creating a case of evaluating the safety effects for the non-ASCS intersections for both EB and FB procedure.

3.3. Investigation of variation of ASCS safety effects

ASCS safety effects could vary with different intersection features. The evaluation results of the safety effectiveness of ASCS are analyzed based on different AADT groups, geometric features, and speed limits of intersections. The evaluation results are aggregated by three groups of AADT at major roads: AADT <= 20,000 vehicles/day, 20,000 vehicles/day < AADT <= 50,000 vehicles/day, and AADT > 50,000 vehicles/day. This grouping of AADT is in line with a previous study (Khattak et al., 2019). The evaluation results are aggregated by two groups based on the number of legs at an intersection (i.e., three-legged and four-legged intersections). The evaluation results are aggregated by six groups based on different speed limits at major roads: 30 mph (13.41 m/s), 35 mph (15.65 m/s), 40 mph (17.88 m/s), 45 mph (20.12 m/s), 50 mph (22.35 m/s), and 55 mph (24.59 m/s). A linear regression model is developed to explore the linear relationship between the ASCS safety effects and each variable (i.e., AADT at major or minor roads, speed limits at major or minor roads, the number of legs at an intersection, the number of exclusive left-turn lanes/right-turn lanes/through lanes on major or minor roads, or the number of access points at an intersection) considered in this study.

4. Data description

As shown in Table 1, reference crash data (i.e., no ASCS is installed) are obtained from similar signalized intersections and corridors (e.g., similar roadway geometrics, the location of proximity, and same functional class of corridors) without ASCS at different locations in South Carolina. Crash data from non-ASCS corridors including US 78 in Berkeley, the segment of US 17A without ASCS in Berkeley, US 1 in Lexington, SC 6 in Lexington, the segment of US 29 without ASCS in Greenville, S-311 in Greenville, SC 146 in Greenville, US 17 in Charleston, SC 171 in Charleston, SC 61 in Charleston, and US 17 in Horry are utilized for the reference crash data. Crash data during before period of ASCS corridors are also utilized for the reference crash data to increase the sample size. The sample size of reference crash data is 680 across different years and different signalized intersections. In the EB procedure, the reference crash data are used for developing the EB models first, and then EB models are combined with the crash data from ASCS corridors to predict EB estimates during after period. Different from the utilization of the crash data in the EB procedure, in the FB procedure, the reference crash data and crash data of ASCS corridors are used directly in the FB models since the FB model prediction and safety effect estimation procedure are conducted in a single step. The South Carolina Department of Transportation (SCDOT) provided the authors with crash data from 2011 to 2018. The crash data include attributes including the crash type and AADT at intersections (major and minor streets). The following roadway geometric features are also collected from Google Earth: (a) the number of exclusive left-turn lanes, right-turn lanes and through lanes on major or minor streets, and (b) the number of access points within the influence area of an intersection. In terms of crash type, crash data are aggregated in four categories: total crashes, F + I crashes, rear-end crashes, and angle crashes. In this paper, intersection crashes are investigated for evaluating the ASCS safety effect. According to SCDOT's strategy, intersection crashes are those that happened within 0.05 miles (80.47 m) of the center of the intersection.

ASCS was not installed in the 24 signalized intersections on US 29 corridor in Greenville, and the corridor could be deemed as a non-ASCS corridor. The crash data of US 29 corridor from 2011 to 2017 during which ASCS was not implemented are used for validating EB and FB models.

Initially, the authors got 13 corridors that have installed ASCS. Original crash data have before period and after period data. The authors only include corridors that have at least two-year after period crash data for this study. ASCS safety effects of six ASCS corridors with a total of 65 signals in South Carolina are evaluated. Only one type of ASCS is investigated in this study.

US 17A in Summerville includes 12 signalized intersections, which have been installed with ASCS since 2015. SC 642 in Charleston consists of 18 signalized intersections, which have been installed with ASCS since 2015. US 52 in Charleston consists of 17 signalized intersections equipped with ASCS since 2016. US 17 in Pawleys Island consists of six signalized intersections equipped with ASCS since 2016. Roper Mt Rd/Garlington Rd in Greenville includes five signalized intersections with ASCS since 2016. N. Lake Drive in Lexington has been implemented with ASCS at seven signalized intersections since 2015. The study crash data pool for safety evaluation excludes crashes that occurred during the ASCS installation year to minimize evaluation bias caused by construction before activating ASCS and driver's adaption to the new driving environment with ASCS.

Few signalized intersections had Flashing Yellow Arrow (FYA) installed during the before or the after period. For this reason, FYA variable was considered as a categorical variable in the initial model to investigate the effects of the number of FYA on the crash frequency at intersections. However, the FYA variable is not significant in the model. Thus, the FYA variable is not included in the model. Offset improvements for left-turn lanes, which have the potential to reduce the number of crashes at signalized intersections, were made on one intersection after the ASCS was installed. Crashes that occurred during the period after offset improvements were made are not included in the analysis. Signal phasing was modified at one of the signals after the ASCS was implemented, so the crashes that occurred during the period after such changes were made are not included in the analysis.

Table 2 shows a summary of descriptive statistics of the geometric features and speed limits at major and minor roads at intersections. The before and after period conditions are similar, in terms of geometric features and speed limits, at major and minor streets at intersections.

Table 3 shows descriptive statistics of the intersection crash frequency (i.e., number of crashes per year) for the before and after period for the ASCS corridors with the maximum number of crashes and minimum number of crashes, respectively. The crash frequency statistics show that crashes are over-dispersed (i.e., variance greater than mean) in the total crash, F + I crash, rear-end crash and angle crash for the ASCS corridors.

Table 1
Crash Data Usage and Resource.

Crash Data Type	Crash Data Resource
Reference Crash Data	Similar signalized corridors without ASCS (US 78 in Berkeley, the segment of US 17A without ASCS in Berkeley, US 1 in Lexington, SC 6 in Lexington, another segment of US 29 without ASCS in Greenville, S-311 in Greenville, SC 146 in Greenville, US 17 in Charleston, SC 171 in Charleston, SC 61 in Charleston, and US 17 in Horry), and ASCS corridors (before period crash data of SC 642, US 52, US 17, Roper Mt Rd/Garlington Rd, N Lake Drive, and US 17A)
Crash Data for Validation of EB and FB Models	Non-ASCS corridor (US 29) with 24 intersections
Crash Data for Safety Evaluation for ASCS Corridors	Six ASCS corridors with 65 intersections (crash data of SC 642, US 52, US 17, Roper Mt Rd/Garlington Rd, N Lake Drive, and US 17A)

5. Validation results of candidate models

This section provides comparison results of the FB and EB models in terms of: (a) the potential bias and variance of prediction and (b) the estimation accuracy of safety effectiveness. Based on the comparison results, this section could guide to select the best model for evaluating the safety effectiveness of ASCS.

Table 2
Descriptive Statistics of Intersection Geometric Features and Speed Limits Data.

Variables	Before Period				After Period			
	Mean	S.D.*	Min	Max	Mean	S.D.*	Min	Max
Number of legs at intersections	3.82	0.38	3	4	3.8	0.4	3	4
Number of through lanes on major streets	5.37	1.44	2	8	5.29	1.28	2	8
Number of the exclusive right-turn lanes on major streets	1.2	0.8	0	2	1.16	0.84	0	2
Number of the exclusive left-turn lanes on major streets	2.28	0.91	0	4	2.22	0.89	0	4
Number of through lanes on minor streets	2.16	1.21	0	5	2.14	1.19	0	5
Number of the exclusive right-turn lanes on minor streets	1.02	0.7	0	2	0.87	0.75	0	2
Number of the exclusive left-turn lanes on minor streets	1.81	0.89	0	4	1.89	0.89	0	4
Number of access points within the influence area of intersection on major streets	3.03	1.75	0	7	3.27	1.8	0	7
Number of access points within the influence area of intersection on minor streets	2.38	1.92	0	7	2.39	1.88	0	7
Speed limit on major streets (mph)	42.64	5	25	55	41.47	5.53	25	55
Speed limit on minor streets (mph)	32.15	4.89	25	50	31.78	4.71	25	50

*S.D.-Standard deviation.

Table 3
Crash Frequency (Number of Crashes per Year) Statistics for ASCS Corridors.

Crash Types	Before period				After period			
	Min	Mean	Max	S.D.*	Min	Mean	Max	S.D.*
US 17A	2011–2014				2016–2018			
Total Crash	5	19.40	52	12.04	7	29.5	86	17.65
F + I	0	4.67	15	3.33	0	5.97	22	4.58
Rear-end	1	9.96	35	8.13	1	14.06	50	10.50
Angle	0	5.88	18	3.76	2	8.06	20	4.16
Roper Mt Rd/Garlington Rd	2011–2015				2017–2018			
Total Crash	0	4.96	23	6.61	0	7.40	28	10.20
F + I	0	0.68	4	1.22	0	0.90	3	1.20
Rear-end	0	3.60	18	4.47	0	5.40	23	7.95
Angle	0	1	8	1.96	0	1.40	7	2.37

*S.D.-Standard deviation.

5.1. Comparison of potential bias and variance of prediction

As shown in Table 4, the FB models have lower RMSE values than that of EB models in all scenarios involving different crash types and predictors. Lower RMSE values indicate lower potential bias and variance of prediction.

5.2. Safety effect estimation comparison

As shown in Fig. 1, Model 6A (AADT + Roadway factor + Year) and Model 6B (AADT + Roadway factor + Year + Spatial effect) have the best estimation because the mean of the crash reduction percentage is quite close to zero (in the "rectangle" box in Fig. 1). This finding indicates that adding the year factor as a covariate into the FB non-spatial model and FB spatial model could improve the accuracy of estimation of the safety effectiveness of ASCS. So safety researchers and practitioners are encouraged to include the year factor in before-and-after evaluation studies.

The difference in the mean of the crash reduction percentage between FB non-spatial models and FB spatial models is small. However, based on the FB spatial model estimation, the spatial effect is statistically significant, which indicates that the spatial effects exist. In addition, DIC is compared between FB non-spatial models and FB spatial models. The difference between the DIC of FB spatial and non-spatial models is more than 10 in all types of models, which indicates that FB spatial models are preferred over the FB non-spatial models. So safety researchers and practitioners are encouraged to include the spatial effects in FB before-and-after evaluation studies.

6. Safety evaluation results

6.1. Corridor-specific evaluation results

Based on the validation results discussed in Section 5, Model 6B that includes AADT, roadway, year factor, and spatial effect, performs best among all models. Six ASCS corridors at different locations in South Carolina are evaluated using Model 6B. Model parameters are not presented in the paper since model parameters for each corridor vary, and presenting model variables will be cumbersome for the paper. Only significant variables of Model 6B for the total crash for SC 642 are shown in Table 5. All variables presented in this table are statistically significant because 95% BCIs do not include zero. A positive sign of an estimate in Table 5 indicates an increase in the number of crashes, while a negative sign of an estimate indicates a reduction in the number of crashes. As presented in Table 5, the variable, the presence of ASCS, is associated with reductions in the number of crashes at intersections. Other variables, year factor, the number of exclusive left-turn lanes on major streets, the number of through lanes/exclusive right-turn lanes/exclusive left-turn lanes on minor streets, the number of access points on major roads, and AADT of major roads and minor roads, are associated with increases in the number of crashes at intersections. The "sigma.spatial effect" variable is statistically significant, indicating that the spatial effects exist on SC 642 and could be captured by a spatial model. The "sigma.random effect" variable is statistically significant, suggesting that the random effect could capture the variations in the crash frequency across intersections and years.

A parameter (the inverse of the square root of the precision parameter indicated in Eq. (22)) of spatial effect estimation is pre-

Table 4
RMSE for EB and FB models.

Model		RMSE			
		Total Crash	F + I	Rear-end	Angle
EB Models	Model 1 (AADT + Annual SPF multipliers)	9.91	5.59	7.07	4.49
	Model 2 (AADT + Road + Annual SPF multipliers)	9.83	5.59	6.92	4.44
	Model 3 (AADT + Road + Year)	9.75	5.54	6.67	4.43
FB Non-spatial Models	Model 4A (AADT)	1.23	1.04	1.31	1.09
	Model 5A (AADT + Road)	1.26	1.01	1.34	1.09
	Model 6A (AADT + Road + Year)	1.15	0.97	1.23	1.01
FB Spatial Models	Model 4B (AADT + Spatial effect)	1.24	0.97	1.30	1.03
	Model 5B (AADT + Road + Spatial effect)	1.31	0.98	1.34	1.05
	Model 6B (AADT + Road + Year + Spatial effect)	1.22	0.91	1.24	0.95

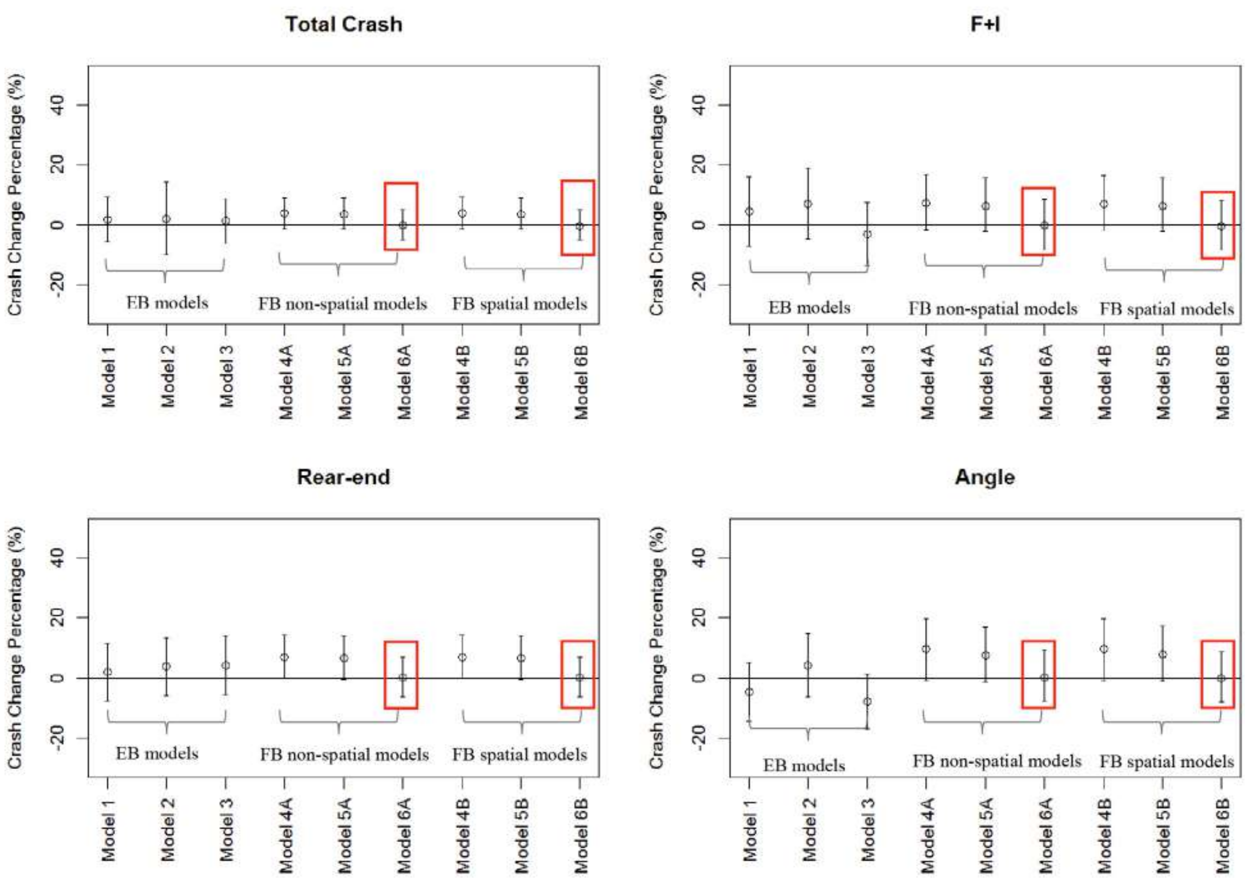


Fig. 1. Crash Change Percentage with 95% CI among EB Models and with 95% BCI among FB Models.

sented in Table 6. The spatial effects are statistically significant for all corridors and crash types since 95% BCIs do not include zero, which indicates that the spatial effects exist on all corridors and could be captured by the spatial model.

Positive signs of values in Table 7 indicate crash increases, while negative signs of values indicate crash reductions. The 95% BCI of each model is shown in the parentheses in Table 7. The ASCS shows crash reductions for the majority of corridors for different crash types.

As shown in Table 7, the highest safety benefits are noted for angle crash for all corridors except US 17A, possibly because the primary objective of the algorithm of ASCS is to minimize total traffic delays of the intersection, which considers the traffic demand from side streets. ASCS potentially decreases the number of angle conflicts.

For rear-end crashes, three corridors (i.e., US 52, N. Lake Drive, and US 17A) shows ASCS increases in rear-end crashes, possibly because ASCS deployed on these corridors tends to achieve balanced service for all vehicle movements, thus minimizing number of stops along corridors (fewer stops may lead to fewer rear-end crashes) tends to be of lower priority than minimizing delay. In addition, the side traffic demand is relatively high among these corridors; thus it may interrupt the major traffic flow.

For US 52, ASCS shows a crash increase in F + I, possibly because US 52 has the highest traffic volume among all corridors, which leads to higher crash severity levels.

Table 5
Model estimates for the total crash evaluation for SC 642.

Variable	Estimate	95% BCI
The presence of ASCS	−0.40	(−0.61, −0.18)
Year factor	0.12	(0.10, 0.15)
The number of exclusive left-turn lanes on major streets	0.07	(0.001, 0.13)
The number of through lanes on minor streets	0.08	(0.03, 0.12)
The number of exclusive right-turn lanes on minor streets	0.22	(0.14, 0.30)
The number of exclusive left-turn lanes on minor streets	0.28	(0.21, 0.35)
The number of access points on major roads	0.06	(0.04, 0.09)
Log (AADT of major roads)	0.75	(0.63, 0.87)
Log (AADT of minor roads)	0.22	(0.18, 0.26)
Intercept	−8.34	(-9.61, −7.08)
sigma.spatial effect[a]	0.65	(0.35, 1.06)
sigma.random effect[b]	0.60	(0.56, 0.64)

a: the inverse of the square root of the precision parameter indicated in Eq. (22).
b: the square root of the variance in Eq. (14).

Table 6
Spatial Effect Estimation for Each Corridor.

Corridor-specific Model	Spatial Effect Estimation (95% BCI)			
	Total Crash	F + I	Rear-end	Angle
SC 642	0.65 (0.35~1.06)	0.30 (0.03~0.89)	0.49 (0.22~0.87)	0.39 (0.09~0.85)
Roper Mt Rd	1.24 (0.15~3.3)	0.67 (0.03~2.81)	0.59 (0.03~2.21)	4.75 (1.15~15.11)
US 17 Pawleys Island	0.28 (0.03~0.92)	0.18 (0.03~0.68)	0.45 (0.05~1.24)	0.18 (0.03~0.67)
US 52	0.31 (0.06~0.71)	0.12 (0.03~0.36)	0.36 (0.06~0.84)	0.48 (0.13~0.96)
N. Lake Drive	0.56 (0.14~1.25)	0.92 (0.19~2.10)	0.27 (0.03~0.81)	0.87 (0.32~1.89)
US 17A	0.33 (0.03~0.84)	0.21 (0.03~0.63)	0.29 (0.03~0.79)	0.45 (0.05~1.01)

Table 7
Corridor-specific Safety Effect Estimation.

Location	Crash Change Percentage (95% BCI)			
	Total Crash	F + I	Rear-end	Angle
SC 642	−32.2%* (−45.0%~−17.4%)	−16.3% (−36.7%~8.5%)	−16.7% (−34.3%~5.1%)	−41.7%* (−55.8%~−24.8%)
Roper Mt Rd	−41.1%* (−64.9%~−8.1%)	−73.7%* (−88.7%~−52.6%)	−3.4% (−45.5%~54.3%)	−92.0%* (−99.4%~−75.3%)
US 17 Pawleys Island	−49.8%* (−66.8%~−27.2%)	−46.7%* (−68.2%~−16.3%)	−39.4%* (−61.1%~−9.8%)	−57.4%* (−73.3%~−35.2%)
US 52	−4.6% (−25.7%~20.8%)	+16.2% (−15.7%~55.9%)	+0.4% (−24.4%~30.5%)	−15.6% (−37.8%~11.8%)
N. Lake Drive	−6.5% (−31.2%~24.4%)	−26.8% (−52.1%~6.4%)	+3.2% (−25.8%~39.5%)	−28.0% (−51.0%~1.8%)
US 17A	+19.7% (−5.7%~19.8%)	−31.8%* (−49.8%~−10.0%)	+17.1% (−9.7%~49.4%)	+10.8% (−15.4%~42.8%)

*: statistically significant in terms of 95% BCI (FB).

6.2. Intersection-specific evaluation results

The safety effectiveness of ASCS is also evaluated for each intersection. As shown in Fig. 2, a negative value means that ASCS reduces crashes. The figure shows that most of the intersections with ASCS show crash reductions for all crash types except the rear-end crash. The ASCS increases rear-end crashes for most of intersections, possibly because ASCS deployed on these intersections tends to achieve balanced service for all vehicle movements, and minimizing the number of stops at intersections (fewer stops may lead to fewer rear-end crashes) tends to be of lower priority than minimizing delay.

The evaluation results are aggregated by three groups of AADT at major roads: AADT less than or equal to 20,000 vehicles/day (sample size = 14), AADT between 20,000 vehicles/day and 50,000 vehicles/day (sample size = 48), and AADT greater than 50,000 vehicles/day (sample size = 3). This grouping of AADT is in line with a previous study (Khattak et al., 2019). As shown in Fig. 3 (b), for F + I cashes, there is a linear relationship between the crash change due to the ASCS and different AADT groups and the linear relationship is inferred based on the regression analysis. Higher AADT decreases ASCS safety benefits in reducing the F + I crashes. The possible reason could be that higher traffic volume may be associated with more severe crashes. As shown in Fig. 3

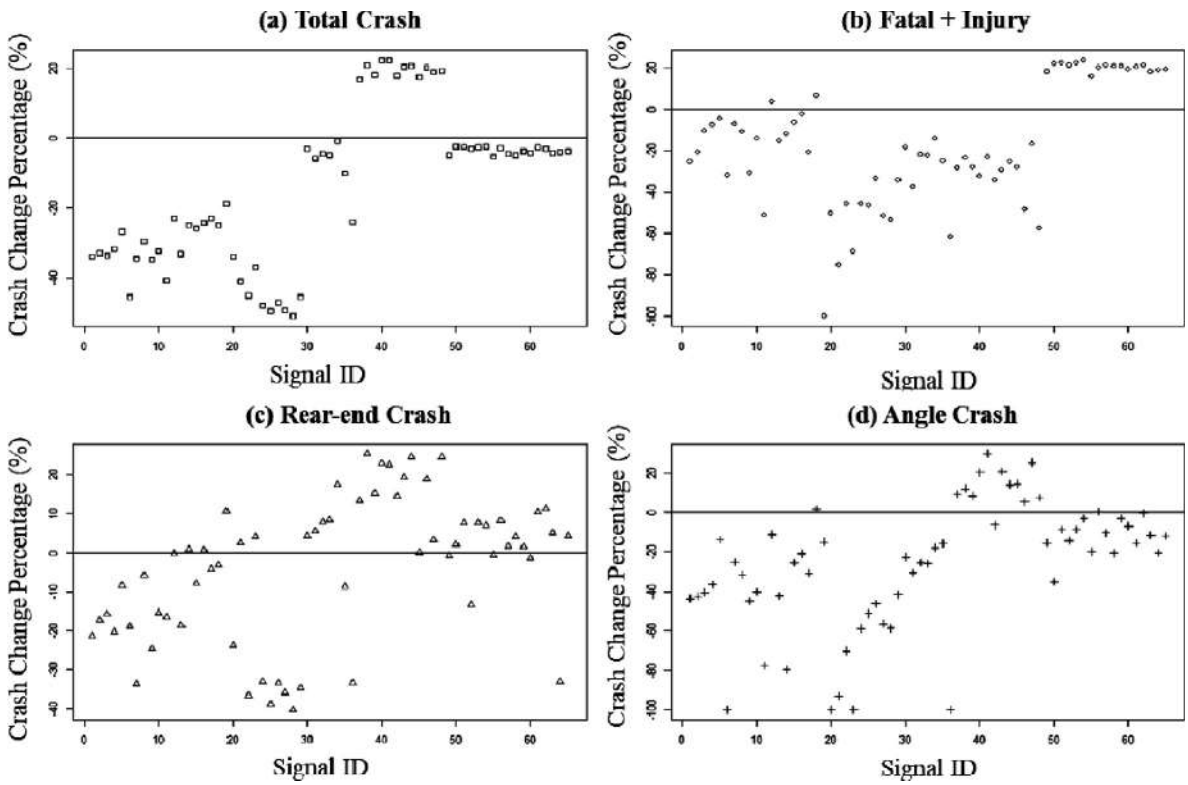


Fig. 2. Percent Change of Crashes due to ASCS at Each Intersection for Different Crash Types.

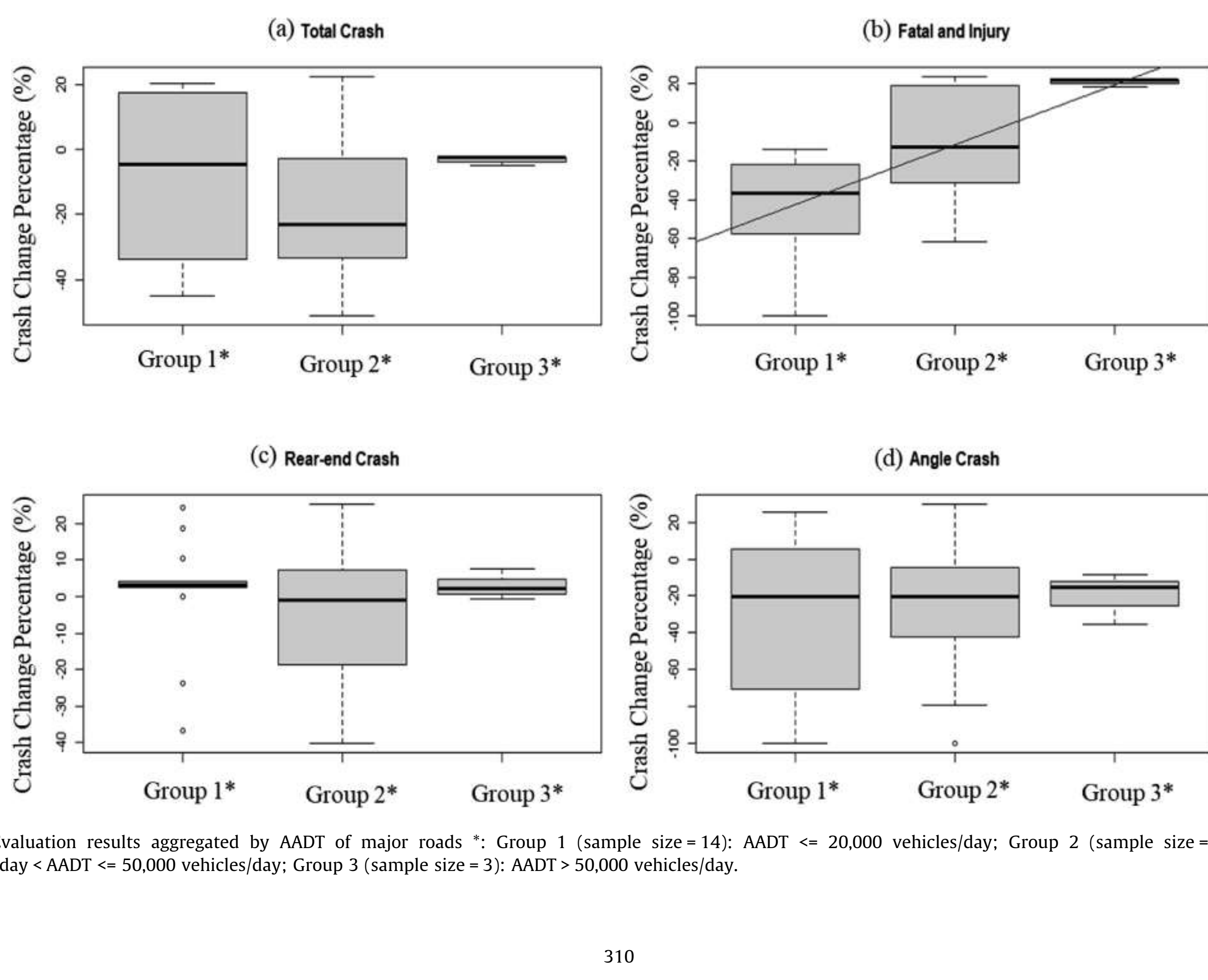


Fig. 3. Evaluation results aggregated by AADT of major roads *: Group 1 (sample size = 14): AADT <= 20,000 vehicles/day; Group 2 (sample size = 48): 20,000 vehicles/day < AADT <= 50,000 vehicles/day; Group 3 (sample size = 3): AADT > 50,000 vehicles/day.

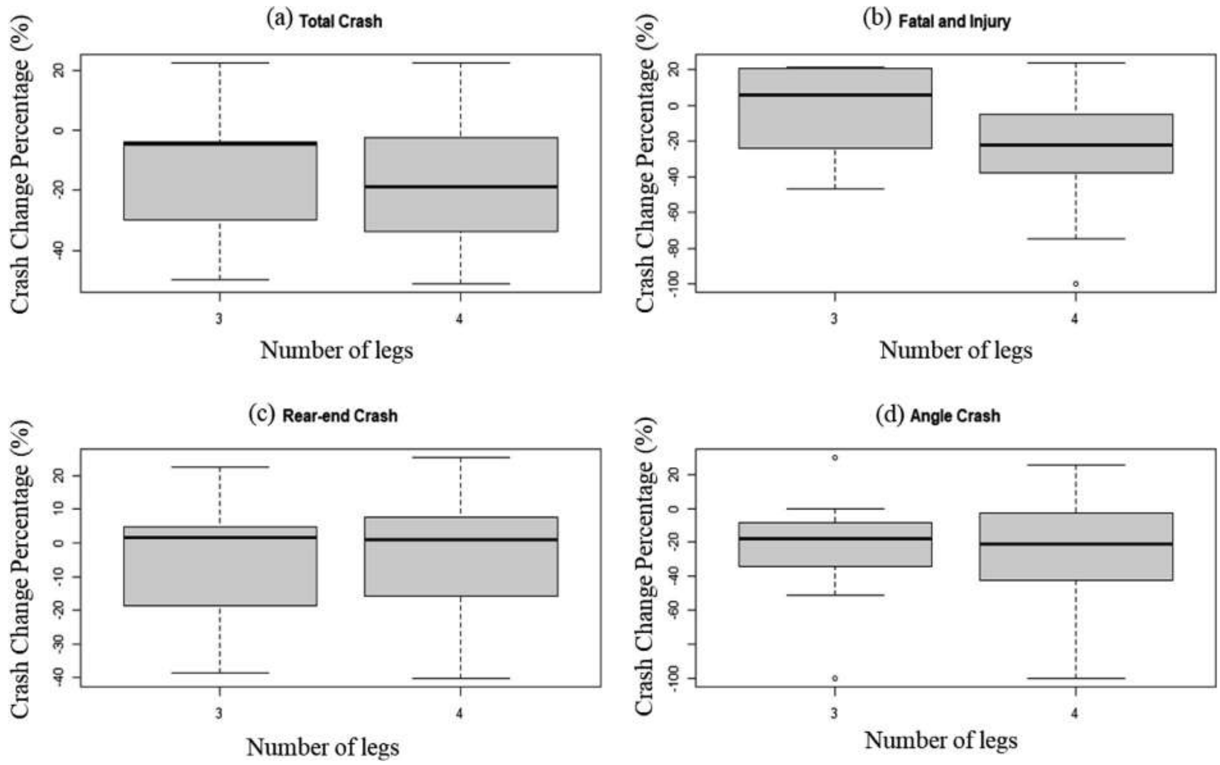


Fig. 4. Evaluation results aggregated by number of legs at an intersection.

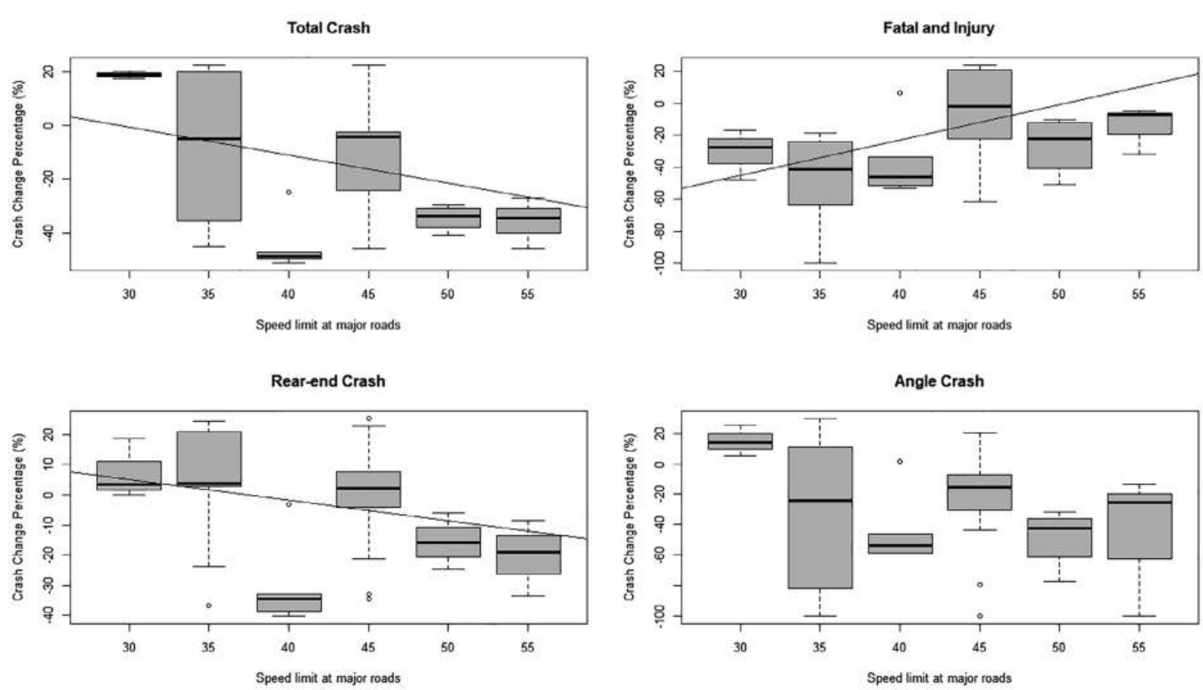


Fig. 5. Evaluation results aggregated by speed limits at major streets**. Sample size for each speed limit: 30 mph: 3; 35 mph: 12; 40 mph: 6; 45 mph: 37; 50 mph: 4, 55 mph: 3.

(a), (c), and (d), for the total crash, rear-end crash, and angle crash, crash changes due to the ASCS are similar for different AADT groups and these crash changes are not statistically different between different AADT groups based on the regression analysis.

The evaluation results are aggregated by two groups based on the number of legs at an intersection, that is, three-legged (sample size = 16) and four-legged intersections (sample size = 49). As shown in Fig. 4 (b), for F + I crashes, the crash reduction due to the ASCS is more considerable in the four-legged intersection compared to the three-legged intersection and the crash reduction due to the ASCS is statistically different between the four-legged intersection and three-legged intersection based on the regression analysis. As shown in Fig. 4 (a), (c), and (d), for the total crash, rear-end crash, and angle crash, crash changes due to the ASCS are similar for four-legged intersections and three-legged intersections and the crash changes due to ASCS are not statistically different between four-legged intersections and three-legged intersections based on the regression analysis.

Additionally, the evaluation results are aggregated by six groups based on different speed limits at major roads: 30 mph (13.41 m/s), 35 mph (15.65 m/s), 40 mph (17.88 m/s), 45 mph (20.12 m/s), 50 mph (22.35 m/s), and 55 mph (24.59 m/s). As shown in Fig. 5 (a) and (c), for the total crash and rear-end crash, there is a linear relationship between the ASCS safety benefits and different speed limits and the linear relationship is inferred based on the regression analysis. The ASCS safety benefit in reducing the total crash and rear-end crash increases as the speed limit increases. As shown in Fig. 5 (b), the ASCS safety benefit in lowering F + I crashes decreases as the speed limit increases and a linear relationship is inferred based on the regression analysis. It is expected that the higher average speed may be associated with higher severe crashes. As shown in Fig. 5 (d), for the angle crash, it is found that there is no linear relationship between the crash change due to the ASCS and different speed limits and it is inferred based on the regression analysis.

A linear regression model is developed to explore the linear relationship between the ASCS safety effects and other variables (i.e., the number of exclusive left-turn lanes/right-turn lanes/through lanes on major or minor streets, and the number of access points at an intersection) considered in this study. Based on our analysis, for F + I crashes, as the number of through lanes on major streets increases, the ASCS safety benefit decreases. More number of through lanes on major streets are associated with higher traffic volume, so the ASCS safety benefit decreases with the increasing traffic volume. For the total crash, rear-end crash, and the angle crash, there is no linear relationship between the safety effectiveness of the ASCS and the number of through lanes on major streets. For the F + I crashes, as the number of access points on minor streets increases, the ASCS safety benefit increases. The possible reason could be that the average speed of the traffic is lower due to the interruption of traffic from/to the access points, so the severe crashes are reduced. For the total crash, rear-end crash, and the angle crash, there is no linear relationship between the safety effectiveness of the ASCS and the number of access points on minor streets.

For all crash types (i.e., total crash, F + I, rear-end crash, and angle crash) considered in this paper, based on the regression analysis, there is no linear relationship between the safety effectiveness of ASCS and AADT of minor roads, the number of the exclusive right-turn lanes on major streets, the number of the exclusive left-turn lanes on major streets, the number of through lanes at minor streets, the number of the exclusive right-turn lanes on minor streets, the number of the exclusive left-turn lanes on minor streets, the number of access points on major streets, and the speed limit at minor streets.

7. Conclusions

This paper develops a series of models, including the Poisson-Lognormal models, Poisson-Gamma models, and spatial models that are implemented in the EB and FB before-and-after studies. Different EB and FB models are validated using real-world non-ASCS intersections. The uniqueness of this paper is that it investigates how model variations would affect: (a) potential bias (e.g., bias due to regression-to-the-mean, traffic volume changes, and roadway geometric feature changes) and variance of prediction and (b) estimation accuracy of safety effectiveness. The findings would provide useful guidance for determining appropriate models for before-and-after safety studies. The FB model that accounts for traffic volume, roadway geometric features, year factor, and spatial effects shows the best performance in reducing potential bias and variance of prediction and improving the accuracy of safety effect estimation.

This paper then applies the best FB model to the safety evaluation of ASCS and evaluates the safety effectiveness of ASCS at six corridors with a total of 65 signalized intersections. ASCS shows crash reductions for most of corridors and intersections. It is also found that the safety effectiveness of ASCS varies with different intersection features (i.e., AADT at major streets, number of legs at an intersection, the number of through lanes on major streets, the number of access points on minor streets, and the speed limit at major streets).

Although this paper discusses different explanatory variables such as AADT, roadway geometric features, and year factor, other possible explanatory variables such as weather conditions, socio-economic factors may be accounted for in developing the crash prediction model. Gaussian CAR distribution is used in the spatial model. However, other distributions of spatial models, such as double exponential distribution and multivariable Gaussian distribution, could be implemented in the spatial model. The effect of neighboring weight matrix structures, such as distance-based weights and exponential decay-based weights on spatial models, may be evaluated in future work.

8. Practical applications

The association between ASCS and crash reductions encourages more ASCS deployments. The variation of the safety effectiveness of ASCS with different intersection features provides insights into selecting ASCS deployment sites for reducing crashes.

9. Declarations of Interest

None.

Acknowledgments

The authors acknowledge the South Carolina Department of Transportation (SCDOT), which provided funding (SCDOT Research Project: SPR 737) for this research. The authors also sincerely acknowledge SCDOT SPR 737 Steering Committee for providing valuable guidance to the authors. The contents of this paper reflect the views of the authors who are responsible for the facts and the accuracy of the presented data. The contents do not reflect the official views of SCDOT or FHWA. This paper does not constitute a standard, specification, or regulation.

References

Antonucci, N. D., Hardy, K. K., Slack, K. L., Pfefer, R., & Neuman, T. R. (2004). Guidance for implementation of the AASHTO strategic highway safety plan. Volume 12: A guide for reducing collisions at signalized intersections. In Transportation Research Board of the National Academies, Washington, DC

Barua, S., El-Basyouny, K., & Islam, M. T. (2016). Multivariate random parameters collision count data models with spatial heterogeneity. *Analytic Methods in Accident Research, 9*, 1–15.

Besag, J., York, J., & Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics, 43*(1), 1–20.

Brooks, S. P., Gelman, A., & Brooks, S. P. (1998). General methods for monitoring convergence of iterative simulations general methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics, 7*(4). https://doi.org/10.1080/10618600.1998.10474787.

Bumham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. New York, New York: Spnnger-Veflag.
Cai, Q., Abdel-Aty, M., Lee, J., Wang, L., & Wang, X. (2018). Developing a grouped random parameters multivariate spatial model to explore zonal effects for segment and intersection crash modeling. *Analytic Methods in Accident Research, 19*, 1–15.
Carriquiry, A. L., & Pawlovich, M. (2004). *From empirical Bayes to full Bayes: Methods for analyzing traffic safety data*. Iowa State University.
Dutta, U., Lynch, J., Dara, B., & Bodke, S. (2010). Safety evaluation of the SCATS control system.
Eghtedari, A. G. (2005). *Measuring the benefits of adaptive traffic signal control: Case study of Mill Plain Blvd. Vancouver, Washington*. Portland State University.
Elkins, S., Niehus, G., Tario, J. D., & Litteer, J. (2012). InSync Adaptive Traffic Control System for the Veterans Memorial Hwy Corridor on Long Island, NY.
Elvik, R., Ulstein, H., Wifstad, K., Syrstad, R. S., Seeberg, A. R., Gulbrandsen, M. U., & Welde, M. (2017). An Empirical Bayes before-after evaluation of road safety effects of a new motorway in Norway. *Accident Analysis & Prevention, 108*, 285–296.
Fink, J., Kwigizile, V., & Oh, J.-S. (2016). Quantifying the impact of adaptive traffic control systems on crash frequency and severity: Evidence from Oakland County, Michigan. *Journal of Safety Research, 57*, 1–7. https://doi.org/10.1016/j.jsr.2016.01.001.
Fontaine, M. D., Ma, J., & Hu, J. (2015). Evaluation of the Virginia Department of Transportation Adaptive Signal Control Technology Pilot Project.
Gross, F., Persaud, B. N., & Lyon, C. (2010). A guide to developing quality crash modification factors.
Hauer, E. (1997). *Observational before/after studies in road safety: Estimating the effect of highway and traffic engineering measures on road safety*. Emerald Group Publishing Limited.
Hauer, E., Harwood, D. W., Council, F. M., & Griffith, M. S. (2002). Estimating safety by the empirical Bayes method: A tutorial. *Transportation Research Record, 1784*(1), 126–131.
Hovey, P. W., & Chowdhury, M. A. (2005). *Development of crash reduction factors*. University of Dayton.
Høye, A. (2015). Safety effects of section control – An empirical Bayes evaluation. *Accident Analysis and Prevention, 74*, 169–178. https://doi.org/10.1016/j.aap.2014.10.016.
Jesus, O. J., & Benekohal, R. F. (2019). Safety analysis and crash modification factors of an adaptive signal control technology along a corridor.
Jin, W., Chowdhury, M., Salek, M. S., Khan, S. M., & Gerard, P. (2021). Investigating hierarchical effects of adaptive signal control system on crash severity using random-parameter ordered regression models incorporating observed heterogeneity. *Accident Analysis and Prevention, 150*. https://doi.org/10.1016/j.aap.2020.105895 105895.
Jonathan, A.-V., Wu, K.-F.-K., & Donnell, E. T. (2016). A multivariate spatial crash frequency model for identifying sites with promise based on crash types. *Accident Analysis & Prevention, 87*, 8–16.
Kergaye, C., Stevanovic, A., & Martin, P. T. (2009). Comparison of before-after versus off–on adaptive traffic control evaluations: Park City, Utah, Case Study. *Transportation Research Record, 2128*(1), 192–201.
Khattak, Z. H. (2016). *Evaluating the operational and safety aspects of adaptive traffic control systems in Pennsylvania*. University of Pittsburgh.
Khattak, Z. H., Fontaine, M. D., Smith, B. L., & Ma, J. (2019). Crash severity effects of adaptive signal control technology: An empirical assessment with insights from Pennsylvania and Virginia. *Accident Analysis & Prevention, 124*, 151–162. https://doi.org/10.1016/j.aap.2019.01.008.
Khattak, Z. H., Magalotti, M. J., & Fontaine, M. D. (2020). Operational performance evaluation of adaptive traffic control systems: A Bayesian modeling approach using real-world GPS and private sector PROBE data. *Journal of Intelligent Transportation Systems, 24*(2), 156–170. https://doi.org/10.1080/15472450.2019.1614445.
Khazraee, S. H., Johnson, V., & Lord, D. (2018). Bayesian Poisson hierarchical models for crash data analysis: Investigating the impact of model choice on site-specific predictions. *Accident Analysis & Prevention, 117*, 181–195. https://doi.org/10.1016/j.aap.2018.04.016.
Lan, B., Persaud, B., Lyon, C., & Bhim, R. (2009). Validation of a full Bayes methodology for observational before–after road safety studies and application to evaluation of rural signal conversions. *Accident Analysis & Prevention, 41*(3), 574–580.
Li, J., & Wang, X. (2017). Safety analysis of urban arterials at the meso level. *Accident Analysis & Prevention, 108*, 100–111.
O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity, 41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6.
Persaud, B., Lan, B., Lyon, C., & Bhim, R. (2010). Comparison of empirical Bayes and full Bayes approaches for before–after road safety evaluations. *Accident Analysis & Prevention, 42*(1), 38–43.
Persaud, B., & Lyon, C. (2007). Empirical Bayes before–after safety studies: Lessons learned from two decades of experience and future directions. *Accident Analysis & Prevention, 39*(3), 546–555.
Sacchi, E., & Sayed, T. (2014). Accounting for heterogeneity among treatment sites and time trends in developing crash modification functions. *Accident Analysis & Prevention, 72*, 116–126.
Sacchi, E., Sayed, T., & El-Basyouny, K. (2016). A full Bayes before-after study accounting for temporal and spatial effects: Evaluating the safety impact of new signal installations. *Accident Analysis & Prevention, 94*, 52–58.
So, J., Stevanovic, A., Posadas, E., & Awwad, R. (2014). Field evaluation of a SynchroGreen adaptive signal system. *T&ampDI Congress, 2014*, 388–399. https://doi.org/10.1061/9780784413586.037.
Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society. Series B: Statistical Methodology, 64*(4), 583–616. https://doi.org/10.1111/1467-9868.00353.
Yanmaz-Tuzel, O., & Ozbay, K. (2010). A comparative Full Bayesian before-and-after analysis and application to urban road safety countermeasures in New Jersey. *Accident Analysis & Prevention, 42*(6), 2099–2107.

Mr. Weimin Jin is a Ph.D. candidate at Clemson University. He received his M.Sc. in 2015 from Tongji University in China. His research interest includes data analytics, regression modeling and inference, probability, Bayesian modeling and inference, traffic safety, connected and automated vehicles, intelligent transportation systems, and traffic simulation.

Dr. Mashrur Chowdhury is the Eugene Douglas Mays Professor in the Glenn Department of Civil Engineering at Clemson University. He is the Director of the USDOT Center for Connected Multimodal Mobility (C^2M^2) (a TIER 1 USDOT University Transportation Center). He is Co-Director of the Complex Systems, Data Analytics and Visualization Institute at Clemson University. He is a Fellow of the American Society of Civil Engineers. He is a registered professional engineer in Ohio.

Dr. Sakib Khan is the Assistant Director of Center for Connected Multimodal Mobility at Clemson University. He was a post-doctoral research scholar working at California Partners for Advanced Transportation Technology (PATH) at the University of California, Berkeley. Before joining the California PATH, he worked as a research specialist for the Center for Connected Multimodal Mobility (C^2M^2) at Clemson University. He received his Ph.D. at 2019, and his M.Sc. at 2015 from Clemson University. He has extensive research experience based on the transportation cyber-physical systems, connected and automated vehicle, and data analytics for transportation. He was also a Co-Principal Investigator of a United States Department of Transportation project.

Dr. Patrick Gerard is a Professor in the School of Mathematical and Statistical Sciences at Clemson University. He currently teaches courses in Multivariate Statistics, Environmental Statistics, and Spatial Statistics. His research includes collaborative efforts with faculty and staff in a wide range of disciplines as well as statistical methodological research primarily in the areas of nonparametric function estimation and computational methods in statistics.

Journal of Safety Research 76 (2021) 135–145

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Investigation of accidents involving powered two wheelers and bicycles – A European in-depth study

Laurie Brown [a,*], Andrew Morris [a], Pete Thomas [a], Karthikeyan Ekambaram [a,1], Dimitris Margaritis [b], Ragnhild Davidse [c], Davide Shingo Usami [d], Massimo Robibaro [d], Luca Persia [d], Ilona Buttler [e], Apostolos Ziakopoulos [f], Athanasios Theofilatos [f,2], George Yannis [f], Alain Martin [g], Fallou Wadji [g]

[a] Transport Safety Research Centre, Loughborough Design School, Loughborough University, Epinal Way, Loughborough, Leicestershire LE11 3TU, UK
[b] Hellenic Institute of Transport, Centre for Research and Technology Hellas (CERTH), 6th Km Charilaou - Thermi R, Thermi, Thessaloniki, Macedonia 57001, Greece
[c] SWOV Institute for Road Safety Research, PO Box 93113, 2509 AC Den Haag, The Netherlands
[d] Centro di Ricerca per il Trasporto e la Logistica (CTL), Sapienza Università di Roma, Via Eudossiana, 18, 00184 Roma, Italy
[e] Motor Transport Institute (ITS), 80 Jagiellońska Str., 03-301 Warsaw, Poland
[f] National Technical University of Athens, Department of Transportation Planning and Engineering, 5 Heroon Polytechniou Str., GR-15773 Zografou, Athens, Greece
[g] European Centre for Safety Studies and Risk Analysis (CEESAR), 132 rue des Suisses, 92000 Nanterre, France

ARTICLE INFO

Article history:
Received 15 May 2020
Received in revised form 21 September 2020
Accepted 18 December 2020
Available online 8 January 2021

Keywords:
Vulnerable road user
Motorcycle
Collision investigation
Crash causation
DREAM

ABSTRACT

Introduction: The number of road fatalities have been falling throughout the European Union (EU) over the past 20 years and most Member States have achieved an overall reduction. Research has mainly focused on protecting car occupants, with car occupant fatalities reducing significantly. However, recently there has been a plateauing in fatalities amongst 'Vulnerable Road Users' (VRUs), and in 2016 accidents involving VRUs accounted for nearly half of all EU road deaths. *Method:* The SaferWheels study collected in-depth data on 500 accidents involving Powered Two-Wheelers (PTWs) and bicycles across six European countries. A standard in-depth accident investigation methodology was used by each team. The Driver Reliability and Error Analysis Method (DREAM) was used to systematically classify accident causation factors. *Results:* The most common causal factors related to errors in observation by the PTW/bicycle rider or the driver of the other vehicle, typically called 'looked but failed to see' accidents. Common scenarios involved the other vehicle turning or crossing in front of the PTW/bicycle. A quarter of serious or fatal injuries to PTW riders occurred in accidents where the rider lost control with no other vehicle involvement. *Conclusions:* Highly detailed data have been collected for 500 accidents involving PTWs or bicycles in the EU. These data can be further analyzed by researchers on a case-study basis to gain detailed insights on such accidents. Preliminary analysis suggests that 'looked but failed to see' remains a common cause, and in many cases the actions of the other vehicle were the critical factor, though PTW rider speed or inexperience played a role in some cases. *Practical Applications:* The collected data can be analyzed to better understand the characteristics and causes of accidents involving PTWs and bicycles in the EU. The results can be used to develop policies aimed at reducing road deaths and injuries to VRUs.

1. Introduction

Road safety remains a major societal challenge within the European Union (EU). In 2016, 25,600 people died on the roads of Europe and 1.4 million people were injured (EC, 2018). Although there are variations between Member States, road fatalities have generally been falling throughout the EU until recent times. During the last few decades, measures to improve road accident prevention have predominantly focused on protecting car occupants to good effect as car occupant fatalities reduced by 44% during the period from 2007 to 2016 (EC, 2018).

* Corresponding author at: Loughborough Design School, Loughborough University, Epinal Way, Loughborough, Leicestershire LE11 3TU, UK.
E-mail address: l.a.brown2@lboro.ac.uk (L. Brown).

[1] Present address: Institute for Future Transport and Cities, Coventry University, Priory St, Coventry CV1 5FB, UK.

[2] Present address: Transport and Urban Planning Group, School of Architecture, Building, and Civil Engineering, Loughborough University, Epinal Way, Loughborough, Leicestershire LE11 3TU, UK.

https://doi.org/10.1016/j.jsr.2020.12.015
0022-4375/

However, at the same time the number of fatalities and injuries amongst other categories of road users has not fallen to the same extent, for example cyclist deaths decreased by only 0.4% on average in the EU between 2010 and 2018 (ETSC, 2020). Vulnerable Road Users (VRUs) are a priority and represent a real challenge for researchers working on accident prevention. Accidents involving VRUs comprised approximately 47% of all fatalities in the EU during 2016. Of these, Powered Two-Wheelers (PTWs) comprised 17% and cyclists 8% of the total numbers of fatalities (EC, 2018), though these proportions do vary between different countries.

Powered two-wheeler is the collective term for motorcycles, mopeds, light mopeds (also called mofas) and speed-pedelecs. PTW use has continued to increase over the years, attracting road users for a variety of reasons such as their lower running costs and ability to easily move in and out of congested traffic (Haworth, 2012). However, there are also disadvantages associated with PTWs, for example they are lightweight and can lose control more easily than a car (Van Elslande & Elvik, 2012). Compared with some other vulnerable road users they can travel at high speeds and mix more closely with other traffic, making them one of the most vulnerable groups of road users and road accidents involving PTW riders are a major social concern.

Bicycle riders are also particularly vulnerable as they travel at lower speeds than motorized vehicles, can be difficult to see, and have little protection if they are involved in an accident. Unlike pedestrians with whom they share these characteristics, cyclists are often on the road mixing directly with other traffic with a higher speed differential, giving them an increased risk of being in an accident with a partner of greater mass.

For these reasons, the SaferWheels study aimed to investigate the causes of accidents involving PTWs and bicycles in Europe. An integral part of this study was that in-depth accident data were collected by trained investigators from six European countries using a common methodology.

The primary objectives of the SaferWheels study were: (1) collection of accident data for at least 500 accidents of which approximately 80% would involve PTWs and the remainder bicycles that collided with a motorized vehicle; (2) in-depth investigations to be carried out using a common established set of protocols based on a systemic approach to risk factor identification; and (3) analysis of the collected data to give an indication of the main accident typologies and causation factors.

It is noted here that the current study only investigated bicycle accidents where a motorized vehicle was involved. Bicycle only, bicycle-bicycle, and bicycle-pedestrian accidents were not in the scope of this study, though research suggests such accidents account for a large number of serious and fatal injuries to cyclists that often go unreported (see for example Schepers, Stipdonk, Methorst & Olivier, 2017, Boele-Vos et al., 2017).

Several previous studies have examined the characteristics of motorcyclist safety. MAIDS (Motorcycle Accidents In-Depth Study), reported in ACEM (2009), carried out in-depth investigations of over 900 accidents involving PTWs in five sampling areas in the EU. The study concluded that the main cause of the majority of PTW accidents was rider or driver error, primarily due to driver inattention, temporary view obstructions or low PTW conspicuity (ACEM, 2009). Other studies have also explored factors affecting injury severity. For example, Albalate and Fernandez-Villadangos (2010) identified gender, excess speed, road width, and alcohol consumption as factors affecting PTW injury severity. Pai and Saleh (2007) determined that junction accidents resulted in more severe outcomes than those not at junctions, and that riding in dark conditions further increased severity. In a recent study, Theofilatos and Ziakopoulos (2018) found that traffic and speed variations increase PTW injury severity, while increased truck proportions in the traffic mix were found to reduce injury severity.

With respect to bicycle accidents, previous literature has identified some common scenarios and causes. Räsänen and Summala (1998) carried out an in-depth analysis of bicycle accidents and found that poor attention allocation and unjustified expectations of the behavior of others were common causes. They also identified a common scenario involving a car driver turning right and coming into conflict with a cyclist on a cycle track. More recently, Wegman, Zhang, and Dijkstra (2012) have explored methods to increase cycling in a population without also increasing fatalities and suggested a safe system approach would best protect vulnerable road users such as cyclists. Tripodi and Persia (2015) further promoted the use of e-safety applications and Information and Communication Technologies (ICT) in enhancing cyclist safety, as well as highlighting that different European countries have varied attitudes to cyclists and so will need different countermeasures.

2. Methodology

Data for the study were collected from sample regions in six EU countries (Table 1) to give a representative view of accidents in Europe. Together the countries accounted for 57% of PTW and 45% of cyclist fatalities in Europe in 2016 (EC, 2018). The sample regions were chosen to be as representative as possible of each country; the relationship between each sample region and the country's national population is described in more detail in Morris et al. (2018).

The objective of the study was to investigate 500 accidents comprising approximately 80% PTW and 20% bicycle accidents; however the proportions would vary for each sample region in order to be more representative of their own accident populations. Table 1 shows the proportion of bicycle and PTW accidents aimed to be investigated by each team to achieve a representative sample. Due to some difficulties in data collection, which are discussed later, these individual proportions were reviewed regularly during the study and adjusted where needed, with some teams collecting more or less PTW / bicycle accidents than originally planned. The numbers that were achieved in practice are shown in the results section in Table 2.

Table 1
Study sampling areas.

Country	Data collection region	Team proportion PTW accidents	Team proportion bicycle accidents
France	Essonne	88%	12%
Greece	Thessaloniki	96%	4%
Italy	Rome	98%	2%
The Netherlands	The Hague	51%	49%
Poland	Mazowieckie	47%	53%
United Kingdom	Midlands	54%	46%

The aim of the study was to investigate the causes of road accidents involving cyclists and PTWs in Europe, therefore only accidents that involved either a PTW or bicycle (or both) were examined. PTW accidents could either be single vehicle or involve a collision partner, however bicycle accidents were only within the sampling criteria if they were in collision with a motorized vehicle. The exception to this was e-bikes (bicycles that provide electrical support even when the cyclist does not pedal at all) and pedelecs (electrically assisted bicycles in which you have to pedal to get assistance), as these could be classified as motorized in their own right and so were included regardless of whether the accident included another motorized vehicle.

For investigation of accidents the study utilized the methodology defined by the DaCoTA project (Atalar, Talbot & Hill, 2012).

The DaCoTA methodology was chosen because: (a) it is a comprehensive guide to conducting in-depth road accident investigations; (b) it has the capability to describe all involved road users in the accident; (c) it has a manual including examples and recommended applications; and (d) it allows all the investigation teams to use a harmonized methodology and thus make the results comparable.

The DaCoTA investigation methodology specifies two primary approaches to gathering information: 'On-Scene' and 'Retrospective.' In the 'On Scene' approach, investigators were notified of an accident by emergency services and attended the scene at the time to collect data. A 'Retrospective' approach was used when attendance at the accident was not possible. In this approach, the vehicles are examined after the accident (e.g., at recovery yards), the scene revisited, and road users approached for interviews. Accident investigation reports (including scene photos, vehicle examinations, driver/rider interviews etc.) from the emergency services are also obtained wherever possible.

The adapted SaferWheels methodology is described fully in Morris et al. (2018). Data were collected during the period of 2015–2017. Investigated accidents usually involved injury to the PTW or bicycle rider; however, a small sample of non-injury accidents were investigated if there were sufficient data available to form a useful case.

A purposive sampling method was adopted. This was based on the concept of saturation, defined as the point at which the data collection process no longer offers any new or relevant data. Case selection was random in all cases, however, there were limitations as not all accidents could be reached in time to investigate thoroughly. Furthermore, barriers such as data privacy issues, legal investigation, explicit refusal by involved parties, etc. prevented the investigation of some accidents. Due to these challenges, some teams relied on investigations of fatal and more serious accidents conducted by specialist police accident investigators ('retrospective' investigations). This did not reflect the true severity distribution of accidents that occur in those regions but was a result of the challenges of collecting in depth accident data.

2.1. Data specification

Approximately 1,500 variables (or fields) per accident were gathered and were entered into a central database. Data were gathered for each element involved in the accident – for example, if the accident involved both a PTW and a passenger car, data were collected for both vehicles and both drivers. The following list illustrates the categories of variables included in the dataset:

- Accident (e.g., date and time, local environment, light and weather conditions)
- Road (e.g., road type, speed limit, road geometry, roadside furniture)
- Road user (e.g., age, gender, injury severity)
- PTW or bicycle (e.g., make and model, motor displacement, mechanical condition)
- Opponent vehicle(s) (e.g., type, make and model, general condition, safety technologies fitted)
- Causation analysis (e.g., speed, distraction, intoxication)
- Reconstruction analysis
- Injury descriptions (coded using the Abbreviated Injury Scale (AIS))
- Road user interviews

2.2. Accident causation classification

Accident causation analysis was carried out using the Driving Reliability and Error Analysis Method (DREAM). DREAM allows investigators to systematically classify and store accident causation information that has been gathered through in-depth investigations by providing a structured method of establishing the causal factors inherent within each accident into a set of formally defined categories of contributing factors.

DREAM originated from the Cognitive Reliability and Error Analysis Model (CREAM) (Hollnagel, 1998), which was used to analyze accidents in process control domains, becoming DREAM when it was adapted for use in road transport accidents (Ljung, 2002, 2007). Warner, Ljung, Sandin, Johansson, and Björklund (2008) developed DREAM further as part of the EC SafetyNet project, and version 3.2, the latest version, was created during the DaCoTA project where additional variables were added specifically relating to PTW accidents (Ljung et al., 2012).

DREAM 3.2 was selected as the preferred method of causation analysis in this study due to the success of previous application, the rigorously established theoretical background, and the structured approach of establishing accident causation specifically for PTWs (Phan et al., 2010).

3. Results

The 500 investigated accidents resulted in a total of 515 'cases;' some accidents involved a PTW and a bicycle so can be considered either a PTW case and/or a bicycle case for analysis purposes. The distribution of PTW and bicycle cases collected by each team is shown in Table 2. in total 77% (385) of the 500 accidents involved a PTW and 26% (130) involved a bicycle.

Table 2
Distribution of cases collected by each investigation team.

Team	PTW cases	Bicycle cases	Total
France	81 (94%)	5 (6%)	86
Greece	78 (92%)	7 (8%)	85
Italy	71 (95%)	4 (5%)	75
The Netherlands	57 (57%)	43 (43%)	100*
Poland	48 (54%)	41 (46%)	89*
United Kingdom	50 (63%)	30 (38%)	80
Total	**385 (75%)**	**130 (25%)**	**515***

*Greater than the total accident number as some accidents involve PTWs and Bicycles.

For the overview analysis e-bikes and pedelecs were grouped with pedal cycles, since they were a small proportion of the sample (14 cases), and share similar characteristics in terms of being able to use cycle lanes, their visibility/conspicuity, and that they generally travel at lower speeds than PTWs.

The distribution of injury severity for the different vehicle types is shown in the Table 3. Teams used the following injury severity classifications:

- Fatal: Death within 30 days of the road accident.
- Serious: Injured (not killed) and hospitalized for at least 24 h.
- Slight: Injured (not killed) and hospitalized for less than 24 h or not hospitalized.
- Not injured: Participated in the accident though not injured.

This classification was applied rather than the national definitions since definitions may vary between countries. The overall distribution comprises 36% (181) slight injury, 30% (149) serious injury, and 17% (84) fatal injury accidents, with the remainder being no injury or unknown severity.

3.1. Accident scenarios

Analysis of the accident scenario was undertaken to look for trends or patterns. The analysis takes into consideration the num-

Table 3
Maximum injury severity of all accidents, PTW cases, and bicycle cases.

Vehicle Type	Non-Injury	Slight	Serious	Fatal	Unknown	Total
PTW cases	22 (5.7%)	134 (34.8%)	103 (26.8%)	69 (17.9%)	57 (14.8%)	385
Bicycle cases	1 (0.8%)	49 (37.7%)	59 (45.4%)	15 (11.5%)	6 (4.6%)	130
All accidents	23 (4.6%)	181 (36.2%)	149 (29.8%)	84 (16.8%)	63 (12.6%)	500

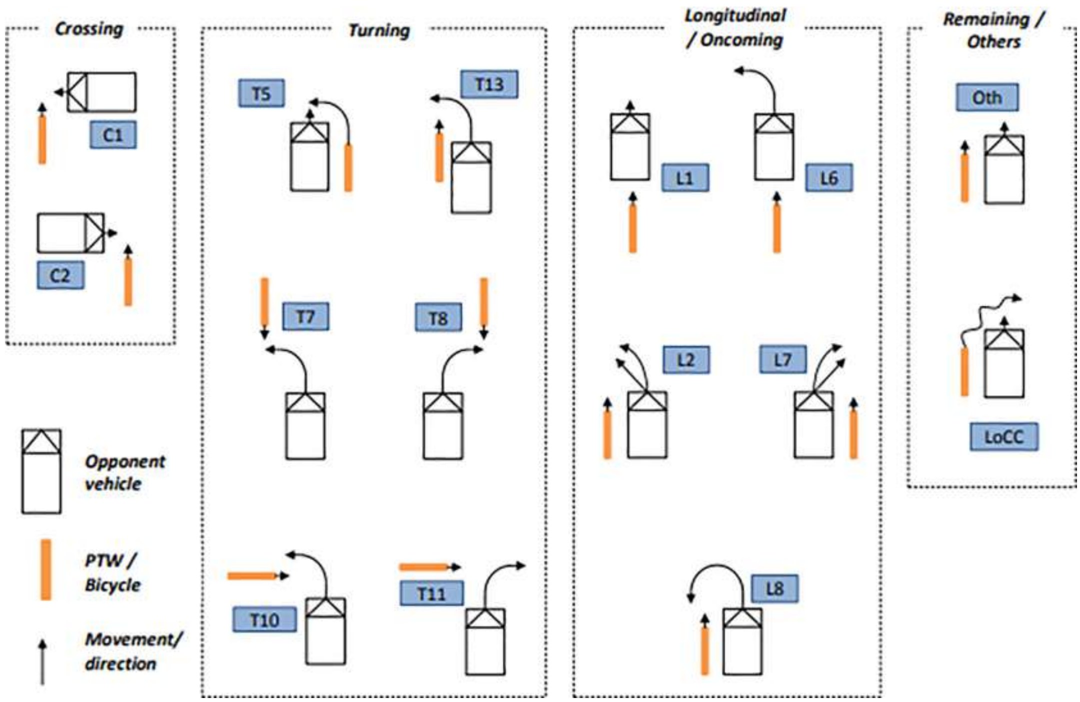


Fig. 1. Grouped accident scenarios for PTW and bicycle accidents (source: Morris et al., 2018).

ber of vehicles/pedestrians involved in the accident, their maneuver, the positions of each road user prior to the accident, and their intended directions. For multi-vehicle accidents, scenario groups were developed for analysis as shown in Fig. 1, which were derived from the 'DaCoTA Accident Type' variable. Further descriptions of these scenarios are included in Appendix A. The main results of the accident scenario analysis are given below.

3.1.1. PTW cases

25% of fatally and seriously injured PTW users were involved in a single vehicle PTW accident. Sixty-four percent of these lost control of their vehicle on a curve. In comparison, only 10% of slight injuries to PTW riders occurred from single vehicle accidents, though it is recognized there may be under-reporting in this area and this figure may not represent the true population.

The three most common accident scenarios for fatally and seriously injured PTW riders involved in a two-vehicle accident were: T7 (16%), C2 (13%) and Loss of Control on a Curve (LoCC – 9%). For slightly injured PTW riders, the two most common accident configurations were T7 (17%) and C1 (16%). The remaining accidents were evenly distributed among the other accident scenarios.

3.1.2. Bicycle cases

The three most common accident scenarios for fatally and seriously injured bicycle riders involved in a two-vehicle accident were C1 (19%), C2 (19%), and T5 (7%). For slightly injured road users involved in a bicycle accident, the three most common accident scenarios differ somewhat; C2 was still the main accident configuration (18%), but the next most frequent were T8 (9%) and T11 (9%).

3.2. Road and environment characteristics

Both PTW and bicycle accidents tended to occur during daylight hours (respectively 78% and 81%). Similarly, most of the accidents occurred under fine dry conditions, with rain, snow, or fog being present in less than 10% of cases. Most of the accidents occurred on urban roads (78% of PTW cases and 83% of bicycle cases), and within a speed limit of 50 km/h or less (79% of all accidents).

Regarding junction-related accidents; 52% of PTW cases and 43% of bicycle cases did not occur at or within 20 m of junctions. When considering all 500 accidents together, 50% occurred at junctions, which was most frequently at a T or Y junction (23%), or crossroads (21%).

3.3. Vehicle characteristics

The sample contains 393 PTW investigations from 385 PTW 'cases' as some accidents involved multiple PTWs. The most common PTW types examined were scooters (47%), followed by road race replicas (19%), standard street bikes (13%), and commuter bikes (7%). The distribution of PTW motor displacement (engine power) is shown in Table 4. Half the sample were lower powered PTWs (250CC or less).

Table 4
Distribution of PTW motor displacement ($n = 393$).

PTW motor displacement	Proportion of sample
50 CC or less	19.8% ($n = 78$)
100–250 CC	29.3% ($n = 115$)
251–500 CC	8.1% ($n = 32$)
Over 500 CC	38.7% ($n = 152$)
Unknown	4.1% ($n = 16$)

The overall PTW condition was coded in the data, ranging from excellent to poor. Excellent or good would indicate the vehicle is in a roadworthy condition, with no obvious signs of defect or poor maintenance. In the majority of cases (80%), the condition of the PTW was found to be good or excellent; only 4% of vehicles were considered to be in poor condition, which would indicate an obvious defect. Defects were observed in 5% of vehicles, most commonly the defects related to the tires, wheel, or brake condition. However, these defects were thought to have contributed to the accident in only 2% of the PTW cases.

Regarding bicycles, 132 bicycles were investigated from 130 bicycle 'cases' as some cases involved two bicycles. Of these, 117 were conventional 'pedal' bicycles and 15 were power assisted (pedelecs). Power assisted bicycles were excluded from detailed bicycle analyses as they have subtle but potentially important differences. Mechanical defects in pedal bicycles were generally limited; when found they were most frequently associated with the tire condition, specifically a worn tread on the tire (11–12% of bicycles). The overall condition of the bicycle was described as good or excellent in 72% of cases. In only 1 case were bicycle defects thought to contribute to the accident.

3.4. Road user characteristics

The 500 investigated accidents involved 1,012 road users, of which 916 (91%) were drivers or riders of the vehicle ($n = 393$ PTW riders, $n = 132$ bicycle riders, $n = 391$ collision opponents). A further 75 (7%) road users were passengers in vehicles, and 21 (2%) were pedestrians; these are generally excluded from analyses unless stated otherwise. PTW riders were highly likely to be male (90%), and two thirds were aged 18–45 (67%). For bicycle riders the gender difference was not as pronounced (68% male), and over half (54%) were over 45 years old.

While most PTW riders used helmets (81%), a non-negligible percentage did not (15%). For bicyclists, only 32% of riders were wearing a fastened helmet; 45% were not wearing one at all. When reading these figures, it is noted that PTW helmets are required by law in all the data collection countries, with an exception that in the Netherlands this only applies to vehicles with an engine displacement over 50 cc. At the time of data collection, light moped riders in the Netherlands were not required to wear helmets, although new laws are being introduced that will change this. Light moped riders in the Netherlands accounted for over half (58%) of the 15% riders who did not wear helmets. In contrast, bicycle helmets are not required by law in any of these countries (apart from in France where they are mandatory only for children under 12 years old), which may in part explain the lower usage observed.

Headlights were used by the majority of PTW riders (72%). However, only 20% of bicycle riders used lights; a further 22% had lights fitted that were not being used and 36% had no lights fitted at all. Reflective and high conspicuity clothing was not often worn by either PTW riders (13%) or bicycle riders (20%). For both headlights and reflective clothing, it should be noted that the figures do not consider the daylight conditions at the time, and the majority of accidents occurred during daylight hours.

3.5. Contributory factors

Contributory factors in more common terminology were derived from the DREAM analyses, which use more specialist terms (e.g., 'attention allocation' became 'distraction'). Through DREAM and other variables in the database nearly 100 possible contributory factors or subfactors were able to be assigned to any given road user. Analyses were carried out for drivers, riders and pedestrians, but not for passengers as they are not in control of the vehicle. Multiple factors were assigned to each road user in each case; in total for the 500 accidents over 4000 factors were assigned with an average of 4.4 factors per road user. Table 5 below shows the results of a selection of 'human' factors commonly related to road accident causation, split by road user type. 'OIRUs' refer to 'other interacting road users' (i.e., drivers of cars/trucks/other vehicles in collision with the PTW or bicycle).

It can be seen that intoxication (alcohol and drug involvement), fatigue, heightened emotions or psychological impairments, medical conditions or physical impairment and risk-taking behavior were not found to be major contributing factors of the investigated accidents. Each of these were thought to be a contributing factor for less than 10% of road users.

Distraction was more prevalent. In particular, for over a third (34%) of the other interacting road users, distraction immediately prior to the accident contributed to its occurrence, compared to 10% of PTW riders, and 16% of cyclists. Distraction could be related to objects/people within the vehicle (e.g. talking to passenger, looking at mobile phone), or outside the vehicle (e.g. focused on road signs, a friend walking past).

Furthermore, errors of observation, typically described as 'looked but failed to see' accidents were a major factor, being a contributing factor for over a third of PTW and bicycle riders (respectively 38% and 39%), and two thirds of interacting road users (66%). Sight obstructions (such as other vehicles, vegetation, or roadside furniture) were also a factor for over a quarter (28%) of interacting road users and may have contributed to some of the errors in observation.

Inexperience as a contributing factor was more prevalent among PTW riders than bicycle riders (respectively 14% and 7%). Inexperience was determined in relation to overall riding experience, familiarity with the specific vehicle ridden, or familiarity with the roads being ridden on. Further analysis was done of the inexperienced PTW riders ($n = 53$). Riders with inexperience as a contributing factor were generally younger, with over half (52%) aged under 25. This is compared with 27% of the total PTW rider sample being in the same age category. Inexperienced riders were also relatively more likely to have speed as a contributing factor compared with all riders (31% compared with 21%).

3.5.1. Speed

Excess speed was rarely observed to be a major factor for cyclists or other interacting road users (respectively 7% and 4%). However, for 22% of PTW riders excess speed was a contributing factor in the accident. The PTW riders that were identified as having speed as a contributing factor ($n = 85$) were further analyzed to determine if there are any trends or commonalities within them.

In the majority of these cases the PTW rider was exceeding the speed limit for the road (71 out of 85 riders), but excess speed was also recorded when the speed was judged to be too fast for the road or weather conditions ($n = 5$ riders speed contributed to the accident but not travelling above the speed limit, $n = 9$ riders speed contributed to the accident but speed limit unknown).

As shown in Table 6, the age profile of riders where speed was a contributing factor is younger than the overall sample, indicating younger people have a higher propensity towards risk taking through speeding. Speed is also correlated with increased injury severities, with PTW accidents where speed was a contributing factor leading to a far higher proportion of fatal/serious injury accidents (81% compared with 45% of all accidents) over slight/no injury accidents (12% compared with 41% of all accidents).

Table 5
Distribution of selected contributory factors according to road user type.

Contributory Factor	Value	Road User		
		PTW Riders ($n = 393$)	Bicycle Riders ($n = 132$)	OIRUs ($n = 391$)
Alcohol*	No	86.8%	84.8%	88.0%
	Yes	4.1%	6.1%	1.5%
	Unknown	9.2%	9.1%	10.5%
Drugs	No	90.6%	92.4%	91.0%
	Yes	3.1%	1.5%	0.5%
	Unknown	6.4%	6.1%	8.4%
Excess Speed	No	54.7%	84.8%	88.5%
	Yes	21.6%	6.8%	3.6%
	Unknown	23.7%	8.3%	7.9%
Fatigue	No	94.9%	89.4%	96.4%
	Yes	2.3%	2.3%	3.1%
	Unknown	2.8%	8.3%	0.5%
Distraction	No	87.5%	75.8%	65.5%
	Yes	9.7%	15.9%	34.0%
	Unknown	2.8%	8.3%	0.5%
Emotional/psychological impairment	No	88.3%	85.6%	93.9%
	Yes	8.9%	6.1%	5.6%
	Unknown	2.8%	8.3%	0.5%
Medical conditions/physical impairment	No	96.2%	88.6%	98.2%
	Yes	1.0%	3.0%	1.3%
	Unknown	2.8%	8.3%	0.5%
Risk-taking behaviour**	No	92.1%	86.4%	98.7%
	Yes	5.1%	5.3%	0.8%
	Unknown	2.8%	8.3%	0.5%
Rider inexperience	No	83.7%	84.8%	95.4%
	Yes	13.5%	6.8%	4.1%
	Unknown	2.8%	8.3%	0.5%
Missed/late observations	No	59.5%	52.3%	33.8%
	Yes	37.7%	39.4%	65.7%
	Unknown	2.8%	8.3%	0.5%
Sight obstruction	No	79.6%	76.5%	71.9%
	Yes	17.6%	15.2%	27.6%
	Unknown	2.8%	8.3%	0.5%

*Note 1: For the Netherlands a large proportion of the data for alcohol involvement were coded as 'unknown', as police in the Netherlands do not regularly check for alcohol involvement.
**Note 2: Factors such as alcohol, drugs and speeding, although also could be considered as risk-taking behaviour, are considered separate to this variable.

Table 6
Age and injury severity of PTW riders for which speed was a contributing factor in the accident ($n = 85$) compared with all riders ($n = 393$).

Age	All PTW riders ($n = 393$)	PTW riders with speed as a contributing factor ($n = 85$)
0–17	4.6%	7.1%
18–25	21.9%	29.4%
26–35	25.2%	25.9%
36–45	19.3%	14.1%
46–55	15.3%	15.3%
56–65	8.9%	7.1%
>65	3.8%	1.2%
Unknown	1.0%	0.0%
Injury Severity	**All PTW accidents ($n = 385$)**	**PTW accidents with speed as a contributing factor ($n = 85$)**
Not injured	5.7%	3.5%
Slight	34.8%	8.2%
Serious	26.8%	31.8%
Fatal	17.9%	49.4%
Unknown	14.8%	7.1%

4. Discussion

4.1. Collection of in-depth accident data

The primary outcome of this study was the collection of in-depth investigation data on 500 accidents involving PTWs or bicycles across six European countries. Many past research studies have raised issues concerning better understanding of the causation of accidents involving VRUs such as PTWs and bicycles, however many of these, for example the MAIDS study (ACEM, 2009), were carried out some time ago. More recently the Motorcycle Crash Causation Study (MCCS) (Nazemetz, Bents, Perry, Thor, & Mohamedshah, 2019) carried out in-depth investigations on 351 PTW accidents in the United States, however there is a lack of more recent large scale in-depth research from a European perspective. The value of the current study therefore is that it will enable researchers to gain a more up to date understanding of the nature and causes of PTW and bicycle accidents in Europe.

The objective of the study was to gather PTW and bicycle accident data from in-depth accident investigations, obtain accident causation and medical data for those accidents, and to store the information according to an appropriate and efficient protocol enabling an accident causation-oriented analysis. The study showed that the DaCoTA protocols for in-depth accident investigations were successful in securing relevant highly detailed data for describing the nature and circumstances of PTW bicycle accidents. Further research could compare the methodology of the current study with other in-depth studies, including the MAIDS and MCCS studies which both utilized adapted OECD investigation protocols.

However, the data collection was not without challenges. Although the target of 500 cases was completed within the time frame of the study, a significant amount of resource was required to achieve this, and some adjustments had to be made to individual team targets and methods of data collection. Ideally an 'on-scene' investigation approach would be used for all cases, as this gives the investigating teams the opportunity to collect more data directly themselves according to the established protocols, supplementing it with additional interview/medical/vehicle examination data later (either directly or through the emergency services). In the current study the on-scene method worked well when utilized and provided accurate data collection, however it was found to have a high cost and time resource associated with it. Teams trying to collect data on-scene faced a variety of challenges, such as; long times 'on-shift' waiting for a suitable accident to occur, not being able to secure data sharing agreements with all emergency services, not being able to reach the accident location before some of the involved parties had already left the scene, being refused permission to interview all involved parties or examine their vehicles, etc. These challenges potentially result in a case not being included in the sample if all the core data could not be collected, in addition to time lost waiting for accidents to occur, so the number of cases collected does not always reflect the amount of effort expended.

Many teams had to instead use the 'retrospective' approach in order to reach their target within the timeframe. Although not able to attend the accident when it occurred, the teams found that it was still possible to gain a large amount of in-depth data by combining data from multiple sources such as police investigation reports and medical examination reports, and that some data could still be collected directly by the investigation teams at a time after the accident (e.g., interviews with road users or examinations of vehicles involved). This method also allowed teams to cover a wider sample area than what they could reach directly from their on-scene base within a short time of the accident occurring, and so increased the sample pool. The drawbacks to this approach include reduction in the amount of data collected, preventing for example a full reconstruction in some cases, and also that investigators often have to rely on interpreting second hand information, which may not have been collected with the same purpose in mind. Overall, however, the retrospective method was found to be more cost-effective, enables better planning and use of staff resources, and does not significantly reduce the quality of the collected data; therefore future studies should consider this method as a good alternative to collecting data directly at the scene if that is not possible.

Finally, issues relating to data protection and privacy, as well as variations in methods or terminology between countries (or different regions within countries) did pose a challenge in collecting harmonized data. The current study highlighted the importance of regular communication between teams through the data collection process to ensure a common understanding was used. Researchers using national datasets to compare accident circumstances between countries face challenges such as incompatible data, missing variables, or unclear definitions. Using a common methodology, in this case the adapted DaCoTA protocols, achieved the aim of generating comparable data between teams, and the output dataset will be valuable in future research to gain insights both within and between countries.

4.2. Data sample

The accident characteristics in the collected sample were in line with those seen in previous in-depth PTW studies such as the MAIDS and MCCS studies, and with similar research on both PTWs and bicycles (e.g., Piantini et al., 2016; Beck et al., 2016). Accidents primarily occurred in urban areas, during daylight hours, and not in adverse weather. Just under half of the accidents investigated occurred at junctions, which is in line with the results reported in the MAIDS study. Compared to the United States, the MCCS study found similar results for single vehicle accidents but reported that over three quarters of multi-vehicle accidents occurred at intersections. Relatively more bicycle accidents occurred at junctions when compared with PTWs, however this could be a function of the sample inclusion criteria as only multi-vehicle bicycle accidents were included, whereas PTWs could be involved in a single-vehicle accident, which often occur outside of junctions.

For both PTWs and bicycles, the characteristics are reported alone and do not consider any exposure data, therefore the results are given solely to describe the sample and do not imply any specific relative risk. Future research could examine this further, considering exposure and comparing with characteristics of all road user types to identify any significant results in the collected data.

4.3. Road user characteristics

In the current study, the PTW rider sample was dominated by males, and over two thirds were aged under 45. This could be explained due to the desires of each age group, as speed, maneuverability and sensation seeking can be said to be the needs of younger people. Conversely, as road-users age, they may seek the comfort of a car, switch to a bicycle or travel on foot, or limit their exposure altogether by taking fewer trips. This was slightly different to the data relating to bicycle riders where two thirds of the sample were male and over half were older than 45. Previous research has found that males are more likely to be involved in a cycling accident (Beck et al., 2016), though this is possibly because of greater use of cyclists by males versus females.

In the accidents investigated, most PTW riders recognized the benefit of helmet use while riding. Haworth and Debnath (2013) found that motorcyclists were more likely to wear a helmet in comparison to cyclists, though this could be related to more legislation being targeted at PTW helmet use. Other research has found that wearing a helmet can reduce injury severity amongst motorcyclists by 70% and reduce the numbers of fatal head injuries by 44% (Elvik, Høye, Vaa, & Sorensen, 2009), which supports the view that continued efforts to improve helmet use by PTW riders will be highly beneficial.

Many of the cyclists investigated in this study did not use a cycle helmet. A recent meta-analysis by Høye (2018) found that in the case of a fall or accident, the use of a bicycle helmet was found to reduce serious head/brain injury by 60% and fatal head/brain injury by 71% on average. However, some studies show adverse effects of bicycle helmets on accident involvement (Robinson, 2006; Phillips, Bjørnskau, Hagmand, & Sagberg, 2011); this is due to 'behavioral adaptation,' as cyclists may feel safer wearing a bicycle helmet and as a result they may show more risky cycling behavior. Other studies indicate that young helmet wearing cyclists take no additional risks (Hagel & Pless, 2006). It is unclear what this could mean for the safety effects of helmet wearing; several studies contradict each other.

4.4. Accident scenarios

A quarter of the serious or fatal PTW cases analyzed involved no other vehicles and two thirds of these were due to the rider losing control on a curve. Loss of control of the PTW was also the third most common accident scenario for multi-vehicle PTW accidents. Combined, these form a large portion of severe outcomes for PTW riders and should be investigated further. Whilst speed is likely to be a factor in a portion of these accidents, it is also recommended to investigate vehicle-based measures to reduce loss of control accidents. See for example Grant et al. (2008), who proposed the implementation of integrated safety systems for a range of PTWs to improve primary safety through handling and stability. Furthermore, Anti-lock Braking Systems (ABS) have been mandatory on European PTWs with engine capacity over 125 cc since 2016 and have been shown to reduce fatal accidents by 31% (Teoh, 2013).

Outside of 'loss of control' accidents, the most common scenarios for multi-vehicle PTW accidents involved another road user turning or crossing in front of the PTW. In most of these cases the PTW rider had the right of way and therefore, although PTW speed was also sometimes a factor, the results show that the actions of the other vehicle drivers are more often the critical factor in the accident than the actions of the PTW rider.

The most common cyclist scenarios also involved other road users crossing in front of them, and as rider speed is highly unlikely to be a factor in bicycle accidents, this suggests failure of other vehicle drivers to either detect them or respond appropriately. More research should be aimed at the other road users involved to better understand why they are committing these right of way violations and to identify how these scenarios can be prevented, for example through the use of in-vehicle intelligent technologies to detect PTWs and bicycles and warn drivers of their presence.

4.5. Contributory factors

The right of way violations may in part be explained by the results seen in the causation analysis. Errors in observation were thought to be a contributory factor for two thirds of the other vehicle drivers analyzed and over a third of PTW and bicycle riders. Interestingly, the MCCS study reported similar results for PTW riders in the United States, but much lower figures for other vehicle drivers (being a cause in less than half of cases). Distraction and sight obstructions were each also prevalent in the current study and are likely to have contributed to the observation errors.

Distraction has long been identified as a common factor in road traffic accidents (e.g., Regan, Lee & Victor, 2013), and it is only expected to increase as both the complexity of the road system increases (e.g., smart motorways, advertising, new vehicle types), and the amount of distractions within vehicles increases (e.g. mobile phones, warnings from driver assistance systems, touchscreen entertainment). The data need to be examined on a case-by-case basis to fully understand the reasons behind the distracted behavior observed, however even from the aggregated analysis it is clear that more measures are needed to combat distraction, potentially through new legislation or targeted awareness campaigns.

However, distraction or obstructions to view did not account for all the observation errors in the analysis, suggesting that 'looked but failed to see' accidents are a large problem for both PTWs and bicycles. Speed of the PTW will have played a role in some cases, as often car drivers can misjudge the speed of an approaching PTW and believe they have time to complete their turn, but a collision occurs when the PTW reaches them sooner than expected (Pai, 2011; Davidse et al., 2019). For non-speed related incidents, particularly those involving crossing or turning scenarios, technology countermeasures might be effective in reducing accidents. Research has shown that vehicle technologies such as advanced forward collision warning would be effective in reducing accidents, including those involving PTWs and bicycles (see e.g. Jermakian, 2011).

Improved conspicuity of riders is also proposed as a countermeasure to this issue. In the investigated accidents use of reflective clothing was low, however De Craen, Doumen, Bos and Van Norden (2011) conclude that it is not so much light or reflective clothing that can increase the visibility of motorcyclists, but particularly the contrast with their environment. Research by Gershon, Ben-Asher, and Shinar (2012) came to a similar conclusion; in an urban environment with a varied and multi-colored background a motorcyclist was more conspicuous in white or reflective clothing, and in a rural setting, where the background mainly consisted of a blue sky, a motorcyclist wearing black was more easily noticed. Clarke, Ward, Bartle and Truman (2004) previously highlighted this problem and calculated that if 'looked but failed to see' errors could be eliminated it could result in a reduction of 25% in the total PTW accident rate. The results of the current study show that this problem is still common over a decade later, so it is clear that more research is required urgently to develop countermeasures to help drivers to recognize PTWs and bicycles and respond appropriately.

Aside from distraction, sight obstructions and errors in observation, the causation analysis did not reveal many other common causes. Intoxication, through alcohol or drugs, and fatigue, whilst traditionally known to be factors in road traffic accidents, did not appear commonly in the current study. Vehicle defects were also not prevalent in the current study and the results show that poor maintenance is not a major cause of PTW or bicycle accidents, being a contributing factor for only 2% of PTW accidents and in only one bicycle accident. Results from both the MAIDS and MCCS studies support that vehicle defects are rarely the primary cause of PTW accidents.

PTW rider inexperience was present in a small but potentially significant amount of cases and was generally associated with younger riders. Lack of experience in driving or riding is a commonly studied factor in road accidents (Groeger, 2006), and accidents can result from a lack of situational awareness, lack of experience in avoiding dangerous situations, or inability to remedy them when they start to occur. Although not analyzed further in the current study, an in-depth review of the cases involving inexperience could give insights into possible countermeasures.

Finally, speed as a contributory factor was analyzed and the results show it was predominately the PTW rider that was speeding, not the bicycle rider or interacting road user. Although only a factor for less than a quarter of cases, preliminary analysis showed that contributory speed was correlated to more severe injury outcomes. The MCCS study similarly reports that excess speed of the PTW was overrepresented in fatal accidents. Much research has been carried out on the benefit of reducing speed on road safety, and the current study supports the view that policies and strategies to reduce speeds would be beneficial in reducing both the number of accidents and their severity. Although the sample was too small to draw statistically significant conclusions, speed being contributory appears to also be correlated with younger inexperienced riders, suggesting that targeted interventions aimed at those groups could be beneficial in reducing accidents.

5. Conclusions

The SaferWheels study collected in-depth investigation data relating to 500 PTW or bicycle accidents within the European Union. These data can be further analyzed by researchers and policy makers to provide insights into how to improve the safety of these vulnerable groups. Accidents involving powered two wheelers and bicycles remain common on European roads and coordinated strategies should be deployed to reduce fatalities and serious injuries. A harmonized dataset containing investigations from six European countries may help towards this, allowing researchers to identify where road safety policies might benefit all member states, and where different countries will need different approaches.

Initial analysis of the 500 investigations reveals that causation factors such as observation errors, distraction, and sight obstructions are particularly prevalent, with 'looked but failed to see' accidents still being a key concern for PTWs and bicycles. Additionally, for PTW riders, there were a small but potentially significant number of cases for which excess speed and/or inexperience was a contributing factor, and these cases could be analyzed further to inform potential countermeasures.

Analysis of the accident scenarios showed that single-vehicle loss of control accidents accounted for a quarter of serious and fatally injured PTW riders, therefore measures to reduce these (through road design, rider behavior or vehicle stability technologies) could result in large benefits for PTW safety. Outside of these, for both PTW and bicycle riders the most common scenarios involved another vehicle crossing or turning in front of them, supporting the view that in many cases, the actions of the PTW or bicycle rider is not the primary factor in the accident.

Whilst the analysis reported here reveals some interesting findings regarding PTW and bicycle accidents, it should be remembered that such findings are based on aggregated analysis of the collective data to look for trends in accident characteristics and causation. Much more can be gained from an evaluation of each individual investigation on a case-study basis to derive more in-depth insight into specific factors that may be relevant to reducing such accidents, as well as how various factors interact with each other to come together and result in an accident.

6. Practical applications

The results of the SaferWheels study have validated the value of a harmonized approach to accident investigation across the European Union, whilst also identifying difficulties in data collection to guide future research methods. The outcome of the study, a dataset of 500 in-depth accident investigations involving PTWs and bicycles, can be analyzed to provide evidence to support policies targeted at reducing road deaths and injuries to vulnerable road users on EU roads.

The SaferWheels dataset is available for analysis upon request from the European Commission.

Acknowledgments

The SaferWheels study was funded by the European Commission under contract number MOVE/C4/2014-661-2. The information and views set out in this study are those of the authors and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this study. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use which may be made of the information contained therein.

The authors would like to offer a sincere thank you to all the road users who have kindly participated in the study. We would also like to thank our colleagues who provided expertise and support in collecting data, and DEKRA and ACEM for their support and advice. Finally, each team would like to thank the police forces, fire and rescue services, ambulance services, hospitals, coroners and government agencies who assisted them in providing access to data for this study.

References

Association of European Motorcycle Manufacturers (ACEM) (2009). MAIDS: In-depth Investigations of Accidents Involving Powered two wheelers, Final report 2.0. ACEM. http://www.maids-study.eu/pdf/MAIDS2.pdf.

Albalate, D., & Fernández-Villadangos, L. (2010). Motorcycle injury severity in Barcelona: The role of vehicle type and congestion. *Traffic Injury Prevention, 11* (6), 623–631. https://doi.org/10.1080/15389588.2010.506932.

Atalar, D., Talbot, R., & Hill, J. (2012). Training Package including training manual and draft protocols, Deliverable 2.3 of the EC FP7 project DaCoTA. European Commission. http://www.dacota-project.eu/Deliverables/DaCoTA-WP2/DaCoTA_D2_3_training%20materials_Final.pdf.

Beck, B., Stevenson, M., Newstead, S., Cameron, P., Judson, R., Edwards, E. R., ... Gabbe, B. (2016). Bicycling crash characteristics: An in-depth crash investigation study. *Accident Analysis & Prevention, 96*, 219–227. https://doi.org/10.1016/j.aap.2016.08.012.

Boele-Vos, M. J., Van Duijvenvoorde, K., Doumen, M. J. A., Duivenvoorden, C. W. A. E., Louwerse, W. J. R., & Davidse, R. J. (2017). Crashes involving cyclists aged 50 and over in the Netherlands: An in-depth study. *Accident Analysis & Prevention, 105*, 4–10. https://doi.org/10.1016/j.aap.2016.07.016.

Clarke, D. D., Ward, P., Bartle, C. & Truman, W., (2004). In-depth study of motorcycle accidents (Road Safety Research Report No. 54). Department for Transport, UK. http://www.righttoride.co.uk/virtuallibrary/ridersafety/indepthstudyofmotorcycleacc2004.pdf.

De Craen, S., Doumen, M., Bos, N. & Van Norden, Y. (2011). The roles of motorcyclists and car drivers in conspicuity-related motorcycle crashes (R-2011-25). SWOV. https://www.swov.nl/en/publication/roles-motorcyclists-and-car-drivers-conspicuity-related-motorcycle-crashes.

Davidse, R. J., van Duijvenvoorde, K., Boele-Vos, M. J., Louwerse, W. J. R., Stelling-Konczak, A., Duivenvoorden, C. W. A. E., & Algera, A. J. (2019). Scenarios of crashes involving light mopeds on urban bicycle paths. *Accident Analysis & Prevention, 129*, 334–341. https://doi.org/10.1016/j.aap.2019.05.016.

Elvik, R., Hoye, A., Vaa, T., & Sorensen, M. (2009). *Handbook of road safety measures* (2nd Edition). Bingley, UK: Emerald Group Publishing.

European Transport Safety Council (ETSC) (2020). How safe is walking and cycling in Europe? (PIN Flash Report 38). Brussels: Dovilé Adminaité-Fodor & Graziella Jost. https://etsc.eu/wp-content/uploads/PIN-Flash-38_FINAL.pdf.

European Commission (EC) (2018). Annual Accident Report 2018. European Commission, Directorate General for Transport. https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/pdf/statistics/dacota/asr2018.pdf.

Gershon, P., Ben-Asher, N., & Shinar, D. (2012). Attention and search conspicuity of motorcycles as a function of their visual context. *Accident Analysis & Prevention, 44*(1), 97–103. https://doi.org/10.1016/j.aap.2010.12.015.

Grant, R., Frampton, R., Peldschus, S., Schuller, E., StClair, V., McCarthy, M., ... & Savino, G. (2008). PISa: powered two-wheeler integrated safety: project objectives, achievements and remaining activities. In International motorcycle conference, 7th, 2008, Essen, Germany (No. 13).

Groeger, J. (2006). Youthfulness, inexperience, and sleep loss: The problems young drivers face and those they pose for us. *Injury Prevention, 12*(Suppl I), 19–24. https://doi.org/10.1136/ip.2006.012070.

Hagel, B. E., & Pless, B. (2006). A critical examination of arguments against bicycle helmet use and legislation. *Accident Analysis & Prevention, 38*(2), 277–278. https://doi.org/10.1016/j.aap.2005.09.004.

Haworth, N. (2012). Powered two wheelers in a changing world – Challenges and opportunities. *Accident Analysis & Prevention, 44*(1), 12–18. https://doi.org/10.1016/j.aap.2010.10.031.

Haworth, N., & Debnath, A. K. (2013). How similar are two-unit bicycle and motorcycle crashes? *Accident Analysis & Prevention, 58*, 15–25. https://doi.org/10.1016/j.aap.2013.04.014.

Hollnagel, E. (1998). *Cognitive reliability and error analysis method: CREAM*. Oxford, UK: Elsevier Science Ltd.

Høye, A. (2018). Bicycle helmets – To wear or not to wear? A meta-analyses of the effects of bicycle helmets on injuries. *Accident Analysis & Prevention, 117*, 85–97. https://doi.org/10.1016/j.aap.2018.03.026.

Jermakian, J. S. (2011). Crash avoidance potential of four passenger vehicle technologies. *Accident Analysis & Prevention, 43*(3), 732–740. https://doi.org/10.1016/j.aap.2010.10.020.

Ljung, M. (2002). *DREAM – Driving Reliability and Error Analysis Method (Master's thesis)*. Linköping University. https://www.diva-portal.org/smash/get/diva2:19361/FULLTEXT01.pdf.

Ljung, M. (2007). Manual for SafetyNet Accident Causation System (SNACS) v1.2. In Reed, S. G. & Morris, A. P. (2008). Glossary of data variables for fatal and accident causation databases - Deliverable 5.5 of the EC FP6 project SafetyNet. http://erso.swov.nl/safetynet/fixed/WP5/D5.5%20Glossary%20of%20Data%20variables%20for%20Fatal%20and%20accident%20causation%20databases.pdf.

Ljung, M., Habibovic, A., Tivesten, S. Sander, J., Bargman, J. & Engstrom, J. (2012). Manual for DREAM v3.2. DaCoTA, EC. https://dacota-investigation-manual.eu/uploads/DREAM_32.pdf.

Morris, A.P., Brown, L.A., Thomas, P., Davidse, R.J., Phan, V., Margaritis, D., Usami, D .... Yannis, G. (2018). SAFERWHEELS study on powered two-wheeler and bicycle accidents in the EU - Final report. European Commission, Directorate General for Mobility and Transport. https://doi.org/10.2832/138260.

Nazemetz, J. W., Bents, F. D., Perry, J. G., Thor, C. P., & Mohamedshah, Y. M. (2019). Motorcycle crash causation study (No. FHWA-HRT-18-064). United States. Federal Highway Administration.
Pai, C. W., & Saleh, W. (2007). An analysis of motorcyclist injury severity under various traffic control measures at three-legged junctions in the UK. *Safety Science, 45*(8), 832–847. https://doi.org/10.1016/j.ssci.2006.08.021.
Pai, C. W. (2011). Motorcycle right-of-way accidents—A literature review. *Accident Analysis & Prevention, 43*(3), 971–982. https://doi.org/10.1016/j.aap.2010.11.024.
Phan, V., Regan, M., Moutreuil, M., Minton, R., Mattsson, M., & Leden, L. (2010). Using the driving reliability and error analysis method (DREAM) to understand powered two-wheeler accident causation. In International conference on safety and mobility of vulnerable road users: Pedestrians, motorcyclists and bicyclists. Jerusalem.
Phillips, R. O., Bjørnskau, T., Hagman, R., & Sagberg, F. (2011). Reduction in car–bicycle conflict at a road–cycle path intersection: Evidence of road user adaptation? *Transportation Research Part F: Traffic Psychology and Behaviour, 14* (2), 87–95. https://doi.org/10.1016/j.trf.2010.11.003.
Piantini, S., Pierini, M., Delogu, M., Baldanzini, N., Franci, A., Mangini, M., & Peris, A. (2016). Injury analysis of powered two-wheeler versus other-vehicle urban accidents. In Proceedings of IRCOBI conference. Malaga, Spain.
Räsänen, M., & Summala, H. (1998). Attention and expectation problems in bicycle–car collisions: An in-depth study. *Accident Analysis & Prevention, 30*(5), 657–666. https://doi.org/10.1016/S0001-4575(98)00007-4.
Regan, M. A., Lee, J. D., & Victor, T. W. (2013). *Driver distraction and inattention: Advances in research and countermeasures (Vol. 1)*. FL, USA: CRC Press.
Robinson, D. L. (2006). No clear evidence from countries that have enforced the wearing of helmets. *British Medical Journal, 332*(7543), 837. https://doi.org/10.1136/bmj.332.7545.837-a.
Schepers, P., Stipdonk, H., Methorst, R., & Olivier, J. (2017). Bicycle fatalities: Trends in crashes with and without motor vehicles in The Netherlands. *Transportation Research Part F: Traffic Psychology and Behaviour, 46*, 491–499. https://doi.org/10.1016/j.trf.2016.05.007.
Teoh, E. R. (2013). *Effects of antilock braking systems on motorcycle fatal crash rates: An update*. USA: Insurance Institute for Highway Safety. https://www.iihs.org/topics/bibliography/ref/2042.
Theofilatos, A., & Ziakopoulos, A. (2018). Examining injury severity of moped and motorcycle occupants with real-time traffic and weather data. *Journal of Transportation Engineering, Part A: Systems, 144*(11), 04018066. https://doi.org/10.1061/JTEPBS.0000193.
Tripodi, A., & Persia, L. (2015). Impact of e-safety applications on cyclists' safety. *International Journal of Injury Control and Safety Promotion, 22*(4), 377–386. https://doi.org/10.1080/17457300.2014.940353.
Van Elslande, P., & Elvick, R. (2012). Powered two-wheelers within the traffic system. *Accident Analysis and Prevention, 49*, 1–4. https://doi.org/10.1016/j.aap.2012.09.007.
Wallén Warner, H., Ljung, M., Sandin, J., Johansson, E., & Björklund, G. (2008). Manual for DREAM 3.0, driving reliability and error analysis method. Deliverable D5. 6 of the EU FP6 project SafetyNet, TREN-04-FP6TRSI2. 395465/506723. Chalmers University of Technology.
Wegman, F., Zhang, F., & Dijkstra, A. (2012). How to make more cycling good for road safety? *Accident Analysis & Prevention, 44*(1), 19–29. https://doi.org/10.1016/j.aap.2010.11.010.

Laurie Brown is a Research Assistant within the Transport Safety Research Centre at Loughborough University and has been active in the field of transport safety for over 7 years. Laurie graduated from the University of York with a degree in Mathematics and Statistics. Currently her research is focussed on safety testing of autonomous vehicles and their interactions with vulnerable road users. She has previously been involved in studies related to the collection and analysis of in-depth road accident data, with a specific focus on vulnerable road users, human factors in transport, and the long-term impacts of road traffic injuries.

Andrew Morris holds a master's degree in Ergonomics and PhD in Mechanical Engineering from the University of Birmingham. He is a Professor of Human Factors in Transport Safety at Loughborough University. In 2009 Andrew was awarded a US Government Special Award of Appreciation in recognition of outstanding leadership and special contributions in the field of motor vehicle safety. Andrew has over 30 years' experience of research in road safety, vehicle safety, accident investigation, crash-injury and driver behaviour. He has published over 200 refereed technical publications, reports and conference papers, and has been involved in successful research grant applications totalling over £9 million.

Pete Thomas is a Professor of Road and Vehicle Safety at Loughborough University and the Academic Lead of the Smart Mobility Living Lab:London CAV test bed, He is a leading figure in the development of road and vehicle safety policy resources and has led the establishment of the European Road Safety Observatory. He has coordinated several large EU research projects into road and vehicle safety and is now examining the societal impact of connected and automated vehicles and mobility services.

Karthik Ekambaram is a Research Fellow at the Institute for Future Transport and Cities within Coventry University, UK. Karthik is currently working on research projects, funded by the UK government, concerning evaluation of Connected and Autonomous Vehicles (CAVs) in real-world (public roads and controlled environment) and virtual settings. Karthik graduated from Anna University, India with an honours degree in Mechanical Engineering and holds post-graduate qualifications from Loughborough University including a Master's degree in Automotive Systems Engineering and PhD degree in Automotive Crashworthiness. Following the completion of his PhD, Karthik was working as a Research Associate at Loughborough University, UK.

Dimitris Margaritis holds a BSc in Automotive Eng, MSc in Environment Protection, MEd in Pedagogical and Technological Education. He is also PhD and Dipl.-Ing. Mechanical Engineer candidate, both at AUTh. He is a Researcher primarily in the area of Traffic Safety and Transport Greening. Since 2006 he works at CERTH/HIT, responsible for the "Institute data analysis, Impact Factors and success Indicators Assessment" Lab. From 1999-2006 he worked for the Crash Safety department at TNO Automotive in The Netherlands. He contributed to the establishment of the first in-depth accident research team in The Netherlands and in Greece.

Ragnhild Davidse is a psychologist and joined SWOV Institute for Road Safety Research in 1996. She is a specialist in the field of older road users and an expert on in-depth accident investigation. She established the SWOV team for in-depth accident investigation in 2008 and has managed several thematic in-depth studies, both for the Dutch Ministry of Infrastructure and Water Management, and for local authorities. Topics included accidents involving run-off-road, light mopeds, bicycles, and mobility scooters. In addition, she has 20 years of experience in research on the safety of older road users and the possibilities to improve their road safety.

Davide Shingo Usami is an Environmental engineer and PhD in Transport Engineering, with more than 10 years of experience in the field of road safety, transport and mobility. He has been involved as an expert in both national and international projects. He is author of publications on journals at national and international level and participates to international Working Groups, such as IRTAD – International Transport Safety Data and Analysis Group.

Massimo Robibaro is an engineer in Computer Science with a PhD in Infrastructure and Transportation with over 7 years of professional experience in design, development, deployment and management of information systems of road safety projects including several years of experience as IT specialist in several projects in EU and developing countries. He collaborated and participated to several international research projects like SAFERWHEELS and DaCoTA for the creation of the database aiming at collecting and analysing in-depth traffic accident data.

Luca Persia, Ph.D. in Transport Planning, is currently Professor and Director of the Research Centre for Transport and Logistics at "La Sapienza" University in Rome. He has studied transport systems for twenty years, and concentrated in studies and research on urban transport systems. He has joined in National and International research projects, in many cases as coordinator, feasibility studies, as well as traffic, road safety and info-mobility plans. He has been supporting Public Administrations in improving road safety in Italy, Brazil, India, Belarus, Cameroon. He is also a reviewer for many international Journals and Magazines.

Ilona Butler is the lead road safety expert for Poland in many international activities. She represents Poland in the European Transport Safety Council activities and has contributed to national road safety strategy. She has been a research partner in many international programmes including the Road Safety Performance Index (PIN) program and the SEC-Safety Belt program coordinated by the ETSC. She has worked on European projects such as BEST (Benchmarking Professional Road Transport Safety, 2001-2003), DRUID (Driving Under the Influence of Drugs, Alcohol and Medicines, 2006-2010). She was also involved in the DaCoTA project (2010-2012) developing a pan-European accident investigation network.

Apostolos Ziakopoulos is a Civil Engineer, MSc, PhD Candidate and Researcher at the Department of Transportation Planning and Engineering at the National Technical University of Athens (NTUA). In 2014, he received a Master of Science – DIC in Transport from Imperial College London and University College London. He has worked as a researcher for the University of Sheffield through the I.A.E.S. T.E. scheme. His main research interests involve road crash analyses, statistical and spatial modelling and traffic engineering. He has participated in 5 research and engineering projects and studies and has published 29 papers in scientific journals and conferences.

Athanasios (Akis) Theofilatos is a Lecturer in Transport Systems at Loughborough University. He holds a Civil Engineering Diploma from National Technical University of Athens (NTUA, 2009), M.Sc. in Transport from Imperial College (2010) and a Ph.D. in Transport Safety from NTUA (2015). His research focuses on road safety, accident analysis, statistical methods, machine learning and stated preference surveys. He has over 8 years' experience in transportation safety and engineering, has participated in 10 research projects and has published 27 journal papers, 2 book chapters and over 40 conference papers. He also serves as a reviewer for over 20 scientific journals.

George Yannis is Professor in Traffic Safety and Management with focus on data management and analysis and is currently Director of the Department of Transportation Planning and Engineering of the School of Civil Engineering at the National Technical University of Athens. For over 30 years, he has contributed extensively in more than 245 research and engineering projects and studies and in several scientific committees of the European Commission and other International Organisations (UNECE, OECD, WHO, World Bank, EIB, CEDR, ERF, IRF, UITP, ETSC, ECTRI, WCTR, TRB). He has published over 560 scientific papers (169 in scientific journals) widely cited worldwide.

Alain Martin is head of branch office of CEESAR (Centre Européen d'Etudes de Sécuritéet d'Analyse des Risques /European Center of Safety Research and Risk analysis) in Essonne. He makes in-depth accident investigation until 2000. For the past 20 years, he has participated in Transport Safety study projects in Europe such as MAIDS (2000-2002), RIDER (2002-2004), Caciauap (2009-2011) and finally Saferwheeks (2015 to 2017). Before this he worked in the commercial field.

Fallou Wadji is an accidentology engineer at CEESAR (Centre Européen d'Etudes de Sécurité et d'Analyse des Risques /European Center of Safety Research and Risk analysis). He makes in-depth accident investigation until 2016. Fallou graduated from the University of Toulouse as an aeronautical mechanics engineer in 2014. He had started working in mechanical calculation and sizing of metal structures between 2014 and 2016.

Appendix A. – Accident Scenario Descriptions

Accident Scenario	Description
C1	• PTW/bicycle driving straight • Opponent vehicle crossing the PTW/bicycle path from the right side
C2	• PTW/bicycle driving straight • Opponent vehicle crossing the PTW/bicycle path from the left side
T5	• PTW/bicycle turning to the left, crossing the (straight) opponent vehicle path • Opponent vehicle is riding straight in the same direction as the heading of the PTW/bicycle before turning
T7	• Opponent vehicle turning to the left, crossing the (straight) PTW/bicycle path • PTW/bicycle coming from the opposite direction, riding straight
T8	• Opponent vehicle turning to the right, crossing the (straight) PTW/bicycle path • PTW/bicycle coming from the opposite direction, riding straight
T10	• Opponent vehicle turning to the left, crossing the (straight) PTW/bicycle path • PTW/bicycle is riding straight, coming from the left side of the opponent vehicle
T11	• Opponent vehicle turning to the right, crossing the (straight) PTW/bicycle path • PTW/bicycle is riding straight, coming from the left side of the opponent vehicle
T13	• Opponent vehicle turning to the left, crossing the (straight) PTW/bicycle path • PTW/bicycle is riding straight in the same direction as the heading of the opponent vehicle before turning
L1	• Opponent vehicle and PTW/bicycle driving in the same direction • PTW/bicycle is riding straight and hit by the opponent vehicle (going straight) from the rear
L2	• Opponent vehicle and PTW/bicycle driving in the same direction • Opponent vehicle is swerving to the left in front of the PTW/bicycle and hit by the PTW/bicycle
L6	• Opponent vehicle and PTW/bicycle driving in the same direction • PTW/bicycle is riding straight and hit by the opponent vehicle (turning left) from the rear
L7	• Opponent vehicle and PTW/bicycle driving in the same direction • Opponent vehicle is swerving to the right in front of the PTW/bicycle and hit by the PTW/bicycle
L8	• Opponent vehicle and PTW/bicycle driving in the same direction • Opponent vehicle is u-turning from the right to the left in front of the PTW/bicycle and hit by the PTW/bicycle
LoCC	• The driver of the PTW/bicycle loses the control of their vehicle, on a curve, and crashes an opponent vehicle
Oth	• All other scenarios that are not covered by any of the previously described scenarios

Journal of Safety Research 76 (2021) 83–89

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Medical referral and license disposition for drivers in Iowa

Jonathan Davis [a,*], Cara Hamann [a,b], Brandon D. Butcher [a,c], Corinne Peek-Asa [a,d]

[a] University of Iowa Injury Prevention Research Center, 2195 Westlawn, Iowa City, IA 52242, USA
[b] Department of Epidemiology, College of Public Health, University of Iowa, 145 North Riverside Drive, Iowa City, IA 52242, USA
[c] Department of Biostatistics, College of Public Health, University of Iowa, 145 North Riverside Drive, Iowa City, IA 52242, USA
[d] Department of Occupational and Environmental Health, College of Public Health, University of Iowa, 145 North Riverside Drive, Iowa City, IA 52242, USA

ARTICLE INFO

Article history:
Received 13 November 2019
Received in revised form 29 July 2020
Accepted 1 December 2020
Available online 19 December 2020

Keywords:
Driver licensing
Fitness to drive
Motor vehicle crash
Older adult
Older driver

ABSTRACT

Introduction: Driver retirement and determination of fitness-to-drive are important aspects of reducing the risk of motor-vehicle collision for an older driver. A lack of information about the review process may lead to poor evaluation of drivers or an increased testing burden to referred drivers. *Methods:* This paper evaluates the license review process for the state of Iowa. We evaluated data from January 2014 to January 2018 and described the source of referral, testing process, and ultimate license disposition. Cox proportional hazards for competing risk were used to determine the risk of having a change in restrictions on the license and the risk of license denial. *Results:* 20,742 individuals were followed through the medical referral process. The most common source of referrals was licensing officials (39.7%). Drivers referred by licensing officials were less likely to be denied their license when compared to drivers from other sources (HR = 0.92 95%CI: 0.87–0.98); however, licensing official referrals were more likely to result in license restrictions compared to other sources (HR = 1.91, 95%CI: 1.82–2.00). Drivers referred by either law enforcement or a physician were more likely to ultimately have their license denied. *Conclusions:* Physician and law enforcement referred the drivers most likely to have their license denied. A smaller proportion of drivers were referred by physicians and law enforcement compared to licensing officials. *Practical Applications:* Licensing agencies should work with physicians and law enforcement to identify drivers who may need a review of their license. Comprehensive tracking of all medical referrals for a driver's license review is important for individual states to understand the burden of their driver referral process and for identifying referral sources with a high proportion of referrals with no licensing change for targeted outreach and education.

1. Introduction

Motor-vehicle collisions are a leading cause of injury and injury-related death in the United States across individuals of driving age (Centers for Disease Control and Prevention, 2017). Decreased physical ability, changes in vision, or medical conditions may result in an increased risk for a crash (Alvarez & Fierro, 2008; Carr, Flood, Steger-May, Schechtman, & Binder, 2006; Carr, Shead, & Storandt, 2005; Emerson et al., 2012; Green, McGwin, & Owsley, 2013). In recognition of this increased risk, each state has provisions for referring drivers for medical review of their fitness-to-drive. Despite this widely accepted means of removing potentially at-risk drivers from the road, there is little existing evidence about the burden, effectiveness, and accuracy of review processes for an individual's fitness-to-drive.

The majority of previous evaluations of the medical review of a license have been limited to referrals made by law enforcement (Lococo, Decina, Branche, & Wagner, 2013; Soderstrom et al., 2010, 2009). A study of 240 drivers over the age of 75 who were referred for medical review in Maryland by law enforcement found that 57% of referred drivers voluntarily gave up their license after referral. Of those who did continue to pursue licensing only half successfully acquired their license at the end of the review process (Soderstrom et al., 2010). Another study of 100 randomly selected police referrals in Virginia found only 12% of drivers were able to continue to drive without changes to their licensure (Lococo et al., 2013). These two studies reveal that a large proportion of drivers referred by police require some change to their driving privileges and that police perform an important role in identifying drivers in need of review. However, drivers referred by law

* Corresponding author at: University of Iowa Injury Prevention Research Center, 2195 Westlawn, Iowa City, IA 52242, USA.

E-mail addresses: jonathan-a-davis@uiowa.edu (J. Davis), cara-hamann@uiowa.edu (C. Hamann), brandon-butcher@uiowa.edu (B.D. Butcher), corinne-peek-asa@uiowa.edu (C. Peek-Asa).

https://doi.org/10.1016/j.jsr.2020.12.001
0022-4375/

enforcement make up only a small proportion of the drivers who are referred for medical review. The evaluation of drivers referred by family members, physicians, or licensing agencies is essential to understanding the medical review process, yet few other studies have examined referrals by all sources and their licensing outcome (Lococo, Sifrit, Stutts, Joyce, & Staplin, 2017; Meuser, Carr, & Ulfarsson, 2009).

In 2014, the Iowa Department of Transportation developed the nation's first integrated system, called the Enhanced Medical Referral and Evaluation Management System (EMREMS), to track medical referrals for fitness-to-drive with driver's licensing and crash outcomes. The objective of this research is to identify which sources of referral for fitness-to-drive are more likely to result in a change in driving privileges or license denial, compared with no changes to licensure status. Additionally, length of referral and screening tests completed by the drivers under review are described to better understand the burden of the review process to these drivers.

2. Methods

2.1. Setting and population

Iowa allows for the voluntary referral or drivers who may need a review of their license based on decreased driving ability. This study includes the analysis of all drivers in Iowa referred for review of their license during January 2014–January 2018. Referral sources were described for all drivers. Drivers were included if they were referred for license review and received a disposition on their license by January 2018.

2.2. Data sources

The Enhanced Medical Referral and Evaluation Management System (EMREMS) is a unique data system in Iowa that tracks medical referrals for fitness-to-drive and the resulting licensure outcomes. A medical referral is a review of a driver's fitness-to-drive based on visual, physical, or cognitive ability. During the referral process, all results of tests evaluating a driver and their licensure outcomes are captured in EMREMS. Data collected during January 2014–January 2018 by the Iowa Department of Transportation and managed with EMREMS were used for this evaluation. The use of EMREMS to track dispositions was fully implemented in 2015 but retrospective case information was available for some drivers prior to 2015. This analysis examined referred drivers' first medical referral. Subsequent referrals were excluded to focus on the initial experience with the referral process. Individuals were excluded if they were reported deceased by the licensing agency or were still under review at the end of January 2018.

The review process can be different for each individual driver. The license review process most frequently includes a request for a vision and/or medical report, as well as, an on-road driving and knowledge test. The tests recorded in EMREMS included the following: Driver Orientation Screen for Cognitive Impairment (DOSCI), Safe Driving BASICS (Brief Auto-Screening Instrument for Cognitive Status), on-road driving test, driving knowledge test, vision screening, and medical report of fitness-to-drive. The DOSCI is a cognitive screening test that entails asking an individual about their name, home address, current location, and time. The DOSCI has been validated with individuals with Alzheimer's and can be used to quickly identify individuals with possible cognitive impairment during a roadside assessment (Hill, Rybar, Stowe, & Jahns, 2016). However, the DOSCI is not designed to specifically identify driving ability. Safe Driving BASICS is an objective, computer-administered battery that combines aspects of visual search, divided attention, visual memory, and visualization of missing information. Individually, these tests have been found to be effective in identifying older individuals with elevated crash risk (Emerson et al., 2012; Hird, Egeto, Fischer, Naglie, & Schweizer, 2016; Jones Ross, Cordazzo, & Scialfa, 2014; Jones Ross, Scialfa, & Cordazzo, 2015; Owsley et al., 1998).

There are several sources of referral into EMREMS, and this analysis included all sources that accounted for more than 1% of referrals. These sources include referrals from the driver themselves (self-referral), law enforcement, crash review, licensing officials, and physicians. A self-referral can result from a driver requesting a review or bringing in a medical or vision report preempting request from the licensing agency itself. This can occur at the request of a physician, but a physician is not considered the source of the review unless they directly request the re-examination of a license. Law enforcement is another source for individuals who commonly submit a request of re-examination to the licensing agency. This can occur after an interaction with a driver. If a crash occurs the circumstances surrounding the crash are reviewed by the licensing agency. At the discretion of the licensing agency and based on circumstances around a crash that suggest diminished driving ability, a review of the license is initiated. Referral from IDOT licensing officials occurs as part of the license renewal if the driver appears to have diminished driving abilities (Iowa Administrative Code 604.50(5)). The official observes the applicant for impairment and asks about changes in medical conditions that may affect the applicant's driving ability. Finally, a review of a license can be set to periodically review a driver's ability. These reviews are designated as recalls. Recalls do not reflect an initial encounter with the license review process. Recalls can be identified in EMREMS as an initial encounter if the previous review occurred before the establishment of the tracking system. Family members and the public can refer drivers, but this occurred at less than 1% in our sample. A request for review is given to the driver in person or mailed to them. Timelines vary for this review, but generally are set to occur within 2 to 3 weeks and can be extended upon request from the driver.

2.3. Statistical analysis

Licensure outcomes included: denial of the license, a change in the restrictions on the license, or no change to the license. A change in restrictions is adding or removing a restriction on the license. The actual change that occurred could not be identified with the available data. It was assumed that these changes are an addition of a restriction as drivers were referred based on possible decreased driving ability. Our primary analysis examined licensure outcomes by the different referral sources. Because recall sources do not represent an initial review and can have resulted from several different sources, the recall source was not analyzed for increased risk of restriction or denial. Cox proportional hazards for competing risk were used to determine the risk of having a change in restrictions on the license and the risk of license denial for individuals, comparing each referral source with drivers referred from other sources (e.g., law enforcement vs. all other sources). Cox proportional hazards for competing risk is an extension of the cox proportional hazard model in which the hazard ratio for receiving a change in license restrictions or having a license denied versus having a license issued without change is estimated. Without the competing risk model, those with a license restriction change would be censored when evaluating the outcome of license denial and the risk would be overestimated. The Cox proportional hazards for competing risk model reduces overestimation of effects by keeping those who experience the competing outcome in the risk set (Austin, Lee, & Fine, 2016; Wolbers, Koller, Witteman, & Steyerberg, 2009). Cox proportional hazards for competing risk has been commonly used in clinical settings when there is a risk for multiple adverse events (Dignam, Zhang, & Kocherginsky, 2012; Gooley, Leisenring, Crowley, & Storer, 1999). The model controlled for age and whether a medical diagnosis was identified during the review. Analysis was repeated excluding drivers who failed to attend or voluntarily surrendered their license. Similarly, analysis was run excluding drivers from a recall source.

Descriptive statistics for sex, age, frequency of diagnoses, length of review, and screening test results were examined. Licensure disposition and length of review were compared across those with and without a medical diagnosis during review of fitness-to-drive using Chi-square and Wilcoxon rank sum tests, respectively. An alpha level of 0.05 was used for all statistical tests.

3. Results

3.1. Demographics

There were 22,238 unique drivers who were referred and received a license disposition for their referral during January 2014 to January 2018. Among these drivers, 311 died during the review and 1,185 had a disposition other than no change, denial, restriction, or surrender of their license (Fig. 1). These cases were excluded. The resulting sample included 20,742 individuals followed through the medical referral process to one of our measured final license dispositions. Demographic information for these individuals is summarized in Table 1. There were about equal numbers of male (53%) and female (47%) drivers referred. The median age of drivers referred was 77 years old (IQR: 54–84) with only 4.5% of drivers under the age of 20.

3.2. Referral and review process

The process by which individuals were referred for medical review and tracked through EMREMS is presented in Fig. 1. The most common source of referral was licensing officials (39.7%), who refer drivers if the license agency employee determines more testing is necessary during license renewal. Individuals evaluated based on a recall of their license (22.5%), as part of a crash review (19.0%), and referred by law enforcement (9.9%) were the other common sources of medical referral. Self-referral (5.5%) and referral by a physician (1.8%) were less common. Together, these sources accounted for 98.3% of all referrals.

During the referral process, slightly less than half (44.8%) of all those referred had a medical condition identified as part of their medical review. The most common medical conditions are presented in Table 1. Vision related conditions were the most common and were comprised of the following diagnoses: macular degeneration ($n = 1030$, 45.9%), cataracts ($n = 970$, 43.2%), glaucoma ($n = 253$, 11.3%), Stargardt macular dystrophy ($n = 34$, 1.5%), hemianopia ($n = 26$, 1.2%), and albinism ($n = 22$, 1.0%). Seizures were the second most common condition identified during review (8.8% of all referrals). The other common conditions were related to consciousness (loss of consciousness and syncope) or cognition (dementia, Alzheimer's, cognitive impairment). Diabetes (2.7%) or stroke (1.7%) made up the remaining types of conditions commonly identified during the review process.

The median time from referral to disposition was 36 days with an interquartile range of 14 to 70 days. Those who had a medical diagnosis had a significantly higher median time for review when compared to those who did not have a medical diagnosis, 42 days versus 33 days (Wilcoxon rank sum P-value <0.0001).

Of drivers with a medical condition, 18.2% had their license denied and 49.9% had a change in their restrictions at the end of the referral process. For those with no medical condition recorded during the review process, 33.7% were ultimately denied their license and 27.6% had a change in their restrictions (Fig. 1). A greater proportion of drivers denied their license without a reported medical condition failed to attend the review (58.1% vs. 31.6%) or voluntarily surrendered their license (22.2% vs. 7.0%) when compared to those with a medical condition identified as part of the referral process. Some drivers who voluntarily surrendered their license or failed to attend part of their review may have a medical condition that was not identified as a result of incomplete testing. When those who failed to attend the required testing or voluntarily surrendered were excluded, 12.2% of drivers with a medical condition were denied their license compared to 8.6% of drivers who did not have a medical condition recorded during the referral process.

Drivers who had their license denied based on a failure to attend differed in their source of referral than drivers who surrendered their license or had their license denied based on a decision

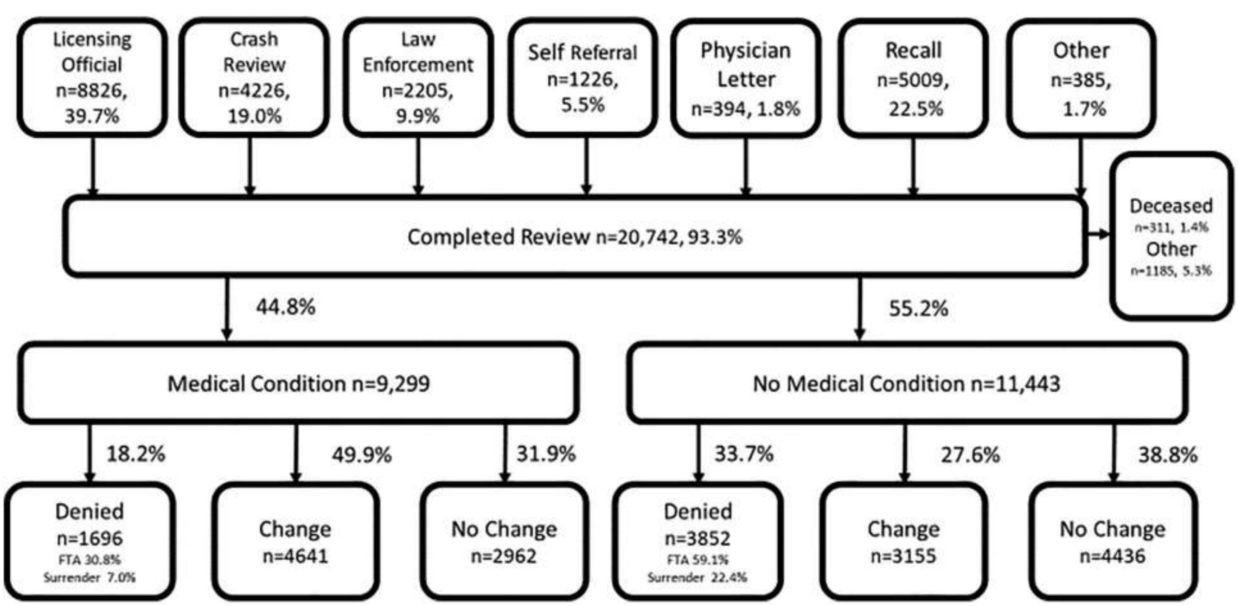


Fig. 1. Flow diagram of drivers referred for licensure review. Abbreviations: FTA, failure to attend.

Table 1
Characteristics of Drivers tracked in EMREMS.

Variable	Level	All		No Change		Restriction**		Denied	
		N = 20742*	%	N = 7398	%	N = 7796	%	N = 5548	%
Sex	Female	9,740	47.0	3458	46.7	3654	46.9	2628	47.4
	Male	10,993	53.0	3939	53.3	4142	53.1	2912	52.6
Age (Years, IQR)	Median	77	(54–84)	76	(52–84)	76	(53–84)	79	(59–86)
Diagnosis†	Any	9,299	44.8	2962	40.0	4641	59.5	1696	30.6
	Vision Related	2,244	10.8	646	8.7	1366	17.5	232	4.2
	Seizure	1,820	8.8	619	8.4	990	12.7	211	3.8
	Loss of Consciousness	719	3.5	221	3.0	306	3.9	192	3.5
	Syncope	266	1.3	60	0.8	115	1.5	91	1.6
	Dementia	334	1.6	17	0.2	82	1.1	235	4.2
	Parkinson's Disease	158	0.8	36	0.5	74	0.9	48	0.9
	Alzheimer's	107	0.5	<10		S		S	
	Cognitive Impairment	104	0.5	10	0.1	40	0.5	54	1.0
	Diabetes	570	2.7	143	1.9	243	3.1	184	3.3
	Stroke	362	1.7	95	1.3	167	2.1	100	1.8
No Diagnosis		11,443	55.2	4436	60.0	3155	40.5	3852	69.4
Time from Referral (Days, IQR)	Median	36	(14–70)	30	(9–56)	32	(10–68)	50	(29–97)

Abbreviations: IQR, Interquartile Range; S, Suppressed – One or more cells for a different group resulted in less than 10 observations.
*Rows may not add to total as a result of missing information.
**Restriction can refer to an adding or removal of a restriction on a license.
† Conditions identified in 100 or more drivers described.

from the licensing agency (data not shown). These drivers were more likely to be referred by law enforcement (21.3% failure to attend drivers vs. 19.2% other denied drivers, p = 0.047) or as part of a crash review (24.6% failure to attend drivers vs. 17.8% other denied drivers, p < 0.001). Drivers who were denied based on a failure to attend were less likely to be referred by a licensing official compared to other denied drivers (30.5% failure to attend drivers vs. 42.7% other denied drivers, p < 0.001). One fifth of the drivers who had a denial based on failure to attend (n = 558, 19.9%) reinitiated the license review process. The additional review resulted in 43.2% (n = 241) of drivers receiving their license with a change to restrictions and 27.4% (n = 153) having their licensed issue without a change to restrictions. Subsequent review resulting in an issuance of a license with or without a change in restrictions accounted for 14.1% (n = 394) of all initial denials based on a failure to attend (n = 2,798).

The hazard ratios for having a change in license restrictions and denial of a license compared to those with no change are presented in Table 2. All models controlled for any diagnosis of a medical condition and age. Drivers had a greater risk of having their license denied if they were referred by law enforcement (HR = 2.57, 95% CI: 2.42–2.74) or a physician (HR = 2.51, 95%CI: 2.17–2.92) compared to drivers from other referral sources while controlling for diagnosis of a medical condition and age. Self-referred drivers were less likely than other drivers to have their license denied (HR = 0.63, 95%CI: 0.53–0.73), but were the most likely to have their license restricted (HR = 1.98, 95%CI: 1.80–2.18). Individuals who were referred from a licensing official were also more likely to have their license restricted than other drivers (HR = 1.91, 95% CI: 1.82–2.00), while drivers referred by a crash review were less likely to have a restriction (HR = 0.42, 95%CI: 0.39–0.46). Drivers referred by a licensing official and crash review were not significantly different from other drivers in regards to denial of a license as a result of the review process.

The crude analysis (not controlling for age and medical diagnosis) was similar in magnitude and direction for most sources (data not shown). However, the crude hazard ratio for denial of drivers referred by crash review compared to other drivers was higher than the adjusted model (HR = 1.32 95%CI: 1.24–1.41).

Analysis was further restricted to drivers who did not voluntarily surrender their license or failed to attend testing. After this restriction the crude and adjusted model for drivers referred as a result of a crash review was similar (HR = 1.09 95%CI: 0.97–1.23). Including only drivers who had a denial as a decision from the DOT resulted in an elevated risk of denial for drivers who were referred from a physician (HR = 4.06 95%CI: 3.29–5.00). This is in part due to the lower number of drivers who failed to attend testing after a physician referral relative to failure to attend denials from other sources. Hazard models were also run, excluding drivers who had recall as their source of referral. No meaningful changes in the results were found, excluding drivers with recall as their source of referral.

3.3. Screening test and evaluation

Table 3 presents the various evaluation tests that drivers completed during review. Screening tests are given at the discretion of the licensing agency. The most common test given to referred drivers was the DOSCI (n = 11,874, 57.2%), followed by the on-road driving test (n = 7,158, 34.5%). The majority of the tests

Table 2
Licensure Outcomes for Drivers tracked in EMREMS.

Source of Referral	Total	Restricted		Denied	
		%	HR*	%	HR*
Licensing Official	8,214	42.1	1.91 (1.82–2.00)	24.7	0.92 (0.87–0.98)
Crash Review	3,917	19.6	0.42 (0.39–0.46)	30.1	1.12 (1.04–1.19)
Law Enforcement	2,012	19.6	0.32 (0.29–0.36)	55.8	2.57 (2.42–2.74)
Self-Referral	1,160	47.7	1.98 (1.80–2.18)	13.9	0.63 (0.53–0.73)
Physician	377	27.3	0.47 (0.39–0.57)	58.4	2.51 (2.17–2.92)

*Model controls for age and whether a medical diagnosis was identified during the review.

Table 3
Screening Tests and Evaluation of Drivers tracked in EMREMS.

Variable	Result	N = 20742	%
On-Road Driving Test	Fail	628	8.8
	Pass	6,530	91.2
	Not Tested	13,584	–
Knowledge Test	Fail	185	4.5
	Pass	3,887	95.5
	Not Tested	16,670	–
Safe Driving BASICS	Fail	611	38.6
	Pass	971	61.4
	Not Tested	19,160	–
DOSCI	Fail	524	4.4
	Pass	11,350	95.6
	Not Tested	8,868	–
Medical Report	Fail	827	13.4
	Pass	5,366	86.6
	Not Tested	14,549	–
Vision Screen	Fail	337	4.9
	Pass	6,550	95.1
	Not Tested	13,855	–

Abbreviations: DOSCI, Driver Orientation Screen for Cognitive Impairment; Safe Driving BASICS, Brief Auto-Screening Instrument for Cognitive Status

resulted in the evaluated driver passing. Only Safe Driving BASICS (38.6% failures) had failure rates above 15%.

4. Discussion

With the use of a state medical referral database, EMREMS, we were able to track medical referrals and final disposition over several years of data. Building on the prior research on referrals from law enforcement, (Lococo et al., 2013; Soderstrom et al., 2009) we evaluated licensure outcomes for several referral sources.

We found that most referrals came as a result of licensing officials at the Department of Transportation. This finding highlights the importance of including other sources of medical referrals when evaluating populations of referred drivers since many referred drivers were not referred from a law enforcement source. Similarly, in a report that compared a random sample of referred drivers from six states (Maine, Ohio, Oregon, Texas, Washington, and Wisconsin) in 2012, non-law enforcement referral sources made up the majority of referrals for all states except Wisconsin (Lococo et al., 2017). The higher representation of law enforcement referrals in Wisconsin was partially due to self-referrals not being tracked in that state.

In the six state report, the primary source of referrals varied across the states, with only Washington having license agency employees as the primary source of referrals, accounting for 18% (Lococo et al., 2017). This finding is only about half the proportion of drivers referred when compared to the results found in EMREMS. The high proportion of drivers referred by a license agency found in our research may be due to different criteria for referral of a driver and subsequent tracking. In Iowa, Department of Transportation employees have discretion to require additional screening of drivers. Drivers asked to complete screening tests, such as the DOSCI or Safe Driving BASICS, are included in the tracking system, possibly explaining the greater representation of these drivers in our study.

A study of Missouri drivers' referral sources found that police officers and license office employees referred the most drivers (Meuser et al., 2009). The study was completed for a time period after the passing of legislation for volunteer referral of drivers with conditions that may affect their fitness-to-drive. Importantly, family members and physicians referred 16% and 20% of drivers, respectively. Referral from family and physicians was much higher than our study despite similar voluntary referral laws. The previous study only included drivers over the age of 50, and many drivers in the study had dementia or a cognitive impairment (45%). Drivers with Alzheimer's, dementia, or a cognitive impairment made up less than 3% of our study population. Even with the differences in age of the populations, this is a great disparity between the frequency of dementia. The difference in underlying population partially explains the higher physician and family reporting sources and is reflective in very few referred drivers keeping their license in the study of Missouri drivers (3.0% of male drivers and 1.8% of female drivers; Meuser et al., 2009). This variability supports the need to track the referral process in states individually.

To our knowledge, this is the first study to compare referral sources and risk of license denial or restriction while controlling for confounders, such as age and medical conditions. In our study, law enforcement and physicians were identified as the sources of referral most likely to result in a denial of a driver's license. However, referrals from physicians were a less common source of a referral than crash reviews, licensing employees, self-referral, and law enforcement. This implies that drivers are unlikely to be referred by a physician, but when they are, they are likely to have their license denied. Physicians may only be referring drivers with severely diminished driving ability. Physicians could be encouraged to lower their threshold for referring drivers for license review since the drivers they currently refer have a high likelihood of license denial. This is similarly true for law enforcement who referred less drivers than the crash review or licensing officials but referred the group of drivers with the second highest risk of denial of their license. This is difficult to implement because of the reliance on personal discretion in making these reviews.

Many of the drivers that had their license denied failed to attend testing. We do not have the data to identify what factors lead to the failure to attend. Iowa DOT requires updated address information for licensing, but this does not guarantee a driver received contact requesting review. About 20% of drivers restarted the review process suggesting some drivers may have been unaware of the review request in addition to those who chose to not attend testing. Further evaluation is needed to understand the motivating reasons for failing to attend testing.

Self-referral was the source most likely to result in a change in the restrictions on a license. This is consistent with the research suggesting that drivers have some ability in determining if their driving skills are diminished (Betz, Carpenter, Genco, & Carr, 2014). Self-referral was rare in our study possibly due to the fear of losing driving privileges. We expect most drivers who self-refer to be encouraged by their physician or a family member, but we are unable to identify this sequence of events in the data. A common source of self-referral is a driver bringing in a medical report or vision report. Again, we expect this is likely at the request of a physician but cannot verify with the available data. The low rate of self-referral may be due to driving changes made by the driver themselves without the intervention of the state licensing authority. The various causes of driving retirement have been summarized previously (Pickard, Tan, Morrow-Howell, & Jung, 2009) as declining health conditions (Dellinger, Sehgal, Sleet, & Barrett-Connor, 2001; Hakamies-Blomqvist & Wahlstrom, 1998; Lyman, McGwin, & Sims, 2001; Ragland, Satariano, & MacLeod, 2004); decreases in visual acuity (Campbell, Bush, & Hale, 1993; West et al., 2003); difficulties performing self-care tasks (Lyman et al., 2001; West et al., 2003); and declining cognitive functions (Freund & Szinovacz, 2002). All of these factors may also result in a lower likelihood to voluntarily seek medical review.

We found that licensing officials were the second most common source of identifying drivers who required a change in their license. License agency employees were by far the most common source of referrals making them the primary identifier of decreased driving abilities. Previous studies have found that in-person license renewal laws reduced the risk of fatal motor-vehicle crash among older drivers regardless of testing required at in-person renewal (Grabowski, Campbell, & Morrisey, 2004; Tefft, 2014). Our results show that licensing agency employees provide a source of identifying the most drivers who require a change to their license. Identifying these drivers and reducing their risk of a future crash through license restriction is one possible explanation for the reduced risk of fatal crash seen in states that require in-person renewal of a license.

During the medical review process the DOSCI was the most frequently used test. This is expected because the DOSCI is a quick screening test and does not require administration by a medical professional (Hill et al., 2016). The tool is convenient but not designed to identify less severe impairments of driving. Another screening test, Safe Driving BASICS was used far less frequently. Safe Driving BASICS may be more appropriate for use in the screening of referred drivers because the individual components have all been effective in identifying diminished driving ability (Emerson et al., 2012; Hird et al., 2016; Jones Ross et al., 2014, 2015; Owsley et al., 1998). This may reduce the need to perform the on-road driving test for some drivers. We found the on-road driving test to be the second most common test administered despite being more intensive than the other screening tests and having a high pass rate. The knowledge test, on-road driving tests, and medical reports are strongly relied on for determining a denial of a license. While the screening tests, such as the Safe Driving BASICS, are not used to determine an outright denial of license, it may be an efficient way to identify drivers who need more extensive review.

While we were able to follow referred drivers from several sources and identify their license disposition, this study does have some limitations. First, we are not able to distinguish what lead to the referral within the reported source. A physician may have encouraged a driver to seek referral or a driver may have requested a physician's evaluation of their driving ability. Either scenario could have initiated the referral process, but only the source listed with the Department of Transportation would have been identified for our study. We were, however, able to generally understand how drivers entered the medical review process. We were also not able to identify if a change in license restrictions resulted in adding or removing restrictions from the available data, only that a change had occurred. Because the drivers were under evaluation of their fitness-to-drive we assume that vast majority of these changes were adding of restrictions.

4.1. Conclusions

This research provides several insights into the medical referral process. Because of their effectiveness in identifying drivers who ultimately have their licenses denied, law enforcement and physicians are valuable partners in identifying diminished driving ability. Future research should evaluate the burden of preventing referral from these sources and provide better guidance to empower these groups to confidently refer drivers. We also found that the screening test, Safe Driving BASICS, was used less frequently than the DOSCI. Based on previous studies, the standardized and objective components of Safe Driving BASICS appear to have broader application in identifying specific impairments that predict crash risk. Drivers have several factors that influence their ability to drive. A more comprehensive analysis of specific medical conditions and how they interact with a screening test will help in the development of guidance needed to more accurately identify drivers no longer fit-to-drive. As EMREMS continues to collect information on referred drivers, these questions can be addressed.

4.2. Practical Applications

Other states should consider tracking their referred drivers in a similar way to evaluate their ability to identify functionally or medically at-risk drivers and make appropriate decisions about their licensure. Because Safe Driving BASICS incorporates several validated measures for identifying those at higher risk of crash involvement, its use should be encouraged. Providing more resources to screen with Safe Driving BASICS may prevent the need for tests that require additional scheduling and employee supervision such as the driving knowledge test or on-road test.

Declaration of Interest

None.

6. Submission declaration

This manuscript has not been published previously and is not under consideration for publication elsewhere.

An initial evaluation of license denial using a subset of this data was presented at the Society for Epidemiologic Research 51st Annual Meeting, Baltimore, MD, 2018. This research includes an additional two years of data and analysis of both license restriction and license denial. We also describe the referral process for the first time, as well as, testing procedures used to evaluate drivers.

Acknowledgements

The authors would like to acknowledge the contributions of the Iowa Department of Transportation and Debra Carney who assisted in gaining access to the data for this project.

Funding

This work was supported by a contract with the Iowa Department of Transportation (DTNH2216RQ-00556, project NTI-12024) and the University of Iowa Injury Prevention Research Center, funded by the CDC National Center for Injury Prevention and Control (R49CE002108).

References

Alvarez, F. J., & Fierro, I. (2008). Older drivers, medical condition, medical impairment and crash risk. *Accident Analysis & Prevention, 40*(1), 55–60. https://doi.org/10.1016/j.aap.2007.04.001.

Austin, P. C., Lee, D. S., & Fine, J. P. (2016). Introduction to the analysis of survival data in the presence of competing risks. *Circulation, 133*(6), 601–609. https://doi.org/10.1161/CIRCULATIONAHA.115.017719.

Betz, M. E., Carpenter, C. R., Genco, E., & Carr, D. B. (2014). Driving self-restriction and age: A study of emergency department patients. *Injury Epidemiology, 1*(1). https://doi.org/10.1186/s40621-014-0018-z.

Campbell, M. K., Bush, T. L., & Hale, W. E. (1993). Medical conditions associated with driving cessation in community-dwelling, ambulatory elders Retrieved from. *Journal of Gerontology, 48*(4), S230–S234 https://www.ncbi.nlm.nih.gov/pubmed/8315247.

Carr, D. B., Flood, K., Steger-May, K., Schechtman, K. B., & Binder, E. F. (2006). Characteristics of frail older adult drivers. *Journal of the American Geriatrics Society, 54*(7), 1125–1129. https://doi.org/10.1111/j.1532-5415.2006.00790.x.

Carr, D. B., Shead, V., & Storandt, M. (2005). Driving cessation in older adults with dementia of the Alzheimer's type Retrieved from. *Gerontologist, 45*(6), 824–827 https://www.ncbi.nlm.nih.gov/pubmed/16326665.

Centers for Disease Control and Prevention (2017). Web-based Injury Statistics Query and Reporting System Retrieved from https://www.cdc.gov/injury/wisqars/index.html.

Dellinger, A. M., Sehgal, M., Sleet, D. A., & Barrett-Connor, E. (2001). Driving cessation: What older former drivers tell us Retrieved from. *Journal of the American Geriatrics Society, 49*(4), 431–435 https://www.ncbi.nlm.nih.gov/pubmed/11347787.

Dignam, J. J., Zhang, Q., & Kocherginsky, M. (2012). the use and interpretation of competing risks regression models. *Clinical Cancer Research, 18*(8), 2301–2308. https://doi.org/10.1158/1078-0432.CCR-11-2097.

Emerson, J. L., Johnson, A. M., Dawson, J. D., Uc, E. Y., Anderson, S. W., & Rizzo, M. (2012). Predictors of driving outcomes in advancing age. *Psychology and Aging, 27*(3), 550–559. https://doi.org/10.1037/a0026359.

Freund, B., & Szinovacz, M. (2002). Effects of cognition on driving involvement among the oldest old: Variations by gender and alternative transportation opportunities Retrieved from. *Gerontologist, 42*(5), 621–633 https://www.ncbi.nlm.nih.gov/pubmed/12351797.

Gooley, T. A., Leisenring, W., Crowley, J., & Storer, B. E. (1999). Estimation of failure probabilities in the presence of competing risks: New representations of old estimators. *Statistics in Medicine, 18*(6), 695–706. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10204198.

Grabowski, D. C., Campbell, C. M., & Morrisey, M. A. (2004). Elderly licensure laws and motor vehicle fatalities. *JAMA, 291*(23), 2840–2846. https://doi.org/10.1001/jama.291.23.2840.

Green, K. A., McGwin, G., Jr., & Owsley, C. (2013). Associations between visual, hearing, and dual sensory impairments and history of motor vehicle collision involvement of older drivers. *Journal of the American Geriatrics Society, 61*(2), 252–257. https://doi.org/10.1111/jgs.12091.

Hakamies-Blomqvist, L., & Wahlstrom, B. (1998). Why do older drivers give up driving? Retrieved from *Accident Analysis and Prevention, 30*(3), 305–312 https://www.ncbi.nlm.nih.gov/pubmed/9663289.

Hill, L. L., Rybar, J., Stowe, J., & Jahns, J. (2016). Development of a curriculum and roadside screening tool for Law enforcement identification of medical impairment in aging drivers. *Injury Epidemiology, 3*(1), 13. https://doi.org/10.1186/s40621-016-0078-3.

Hird, M. A., Egeto, P., Fischer, C. E., Naglie, G., Schweizer, T. A., & Pachana, N. (2016). A systematic review and meta-analysis of on-road simulator and cognitive driving assessment in Alzheimer's disease and mild cognitive impairment. *Journal of Alzheimer's Disease, 53*(2), 713–729. https://doi.org/10.3233/JAD-160276.

Jones Ross, R. W., Cordazzo, S. T. D., & Scialfa, C. T. (2014). Predicting on-road driving performance and safety in healthy older adults. *Journal of Safety Research, 51*, 73–80. https://doi.org/10.1016/j.jsr.2014.09.005.

Jones Ross, R. W., Scialfa, C. T., & Cordazzo, S. T. D. (2015). Predicting on-road driving performance and safety in cognitively impaired older adults. *Journal of the American Geriatrics Society, 63*(11), 2365–2369. https://doi.org/10.1111/jgs.13712.

Lococo, K. H., Decina, L. E., Branche, J., & Wagner, E. M. (2013). Medical review process and license disposition of drivers referred by law enforcement in Virginia. *Journal of Safety Research, 45*, 47–53. https://doi.org/10.1016/j.jsr.2013.01.005.

Lococo, K. H., Sifrit, K. J., Stutts, J., Joyce, J. J., & Staplin, L. (2017). Medical review practices for driver licensing, Volume 2: Case studies of medical referrals and licensing outcomes in six states. Retrieved from Washington, DC.

Lyman, J. M., McGwin, G., Jr., & Sims, R. V. (2001). Factors related to driving difficulty and habits in older drivers Retrieved from. *Accident Analysis and Prevention, 33*(3), 413–421 https://www.ncbi.nlm.nih.gov/pubmed/11235803.

Meuser, T. M., Carr, D. B., & Ulfarsson, G. F. (2009). Motor-vehicle crash history and licensing outcomes for older drivers reported as medically impaired in Missouri. *Accident Analysis & Prevention, 41*(2), 246–252. https://doi.org/10.1016/j.aap.2008.11.003.

Owsley, C., Ball, K., McGwin, G., Jr., Sloane, M. E., Roenker, D. L., White, M. F., & Overley, E. T. (1998). Visual processing impairment and risk of motor vehicle crash among older adults Retrieved from. *JAMA, 279*(14), 1083–1088 https://www.ncbi.nlm.nih.gov/pubmed/9546567.

Pickard, J. G., Tan, J., Morrow-Howell, N., & Jung, Y. (2009). Older drivers retiring from the road: An application of the selection, optimization, and compensation model. *Journal of Human Behavior in the Social Environment, 19*(2), 213–229. https://doi.org/10.1080/10911350802687232.

Ragland, D. R., Satariano, W. A., & MacLeod, K. E. (2004). Reasons given by older people for limitation or avoidance of driving Retrieved from. *Gerontologist, 44*(2), 237–244 https://www.ncbi.nlm.nih.gov/pubmed/15075420.

Soderstrom, C. A., Scottino, M. A., Burch, C. A., Ho, S. M., Kerns, T. J., & Joyce, J. J. (2010). Pursuit of licensure by senior drivers referred by police to a State Licensing Agency's Medical Advisory Board Retrieved from. *Ann Adv Automot Med, 54*, 351–358 https://www.ncbi.nlm.nih.gov/pubmed/21050617.

Soderstrom, C. A., Scottino, M. A., Joyce, J. J., Burch, C., Ho, S. M., & Kerns, T. J. (2009). Police referral of drivers to the Maryland Motor Vehicle Administration's Medical Advisory Board Retrieved from. *Ann Adv Automot Med, 53*, 105–116 https://www.ncbi.nlm.nih.gov/pubmed/20184837.

Tefft, B. C. (2014). Driver license renewal policies and fatal crash involvement rates of older drivers, United States, 1986–2011. *Injury Epidemiology, 1*(1), 25. https://doi.org/10.1186/s40621-014-0025-0.

West, C. G., Gildengorin, G., Haegerstrom-Portnoy, G., Lott, L. A., Schneck, M. E., & Brabyn, J. A. (2003). Vision and driving self-restriction in older adults Retrieved from. *Journal of the American Geriatrics Society, 51*(10), 1348–1355 https://www.ncbi.nlm.nih.gov/pubmed/14511153.

Wolbers, M., Koller, M. T., Witteman, J. C. M., & Steyerberg, E. W. (2009). Prognostic models with competing risks: Methods and application to coronary risk prediction. *Epidemiology, 20*(4), 555–561. https://doi.org/10.1097/EDE.0b013e3181a39056.

Jonathan Davis is a postdoctoral research scholar at the University of Iowa Injury Prevention Research Center. His research investigates older adult safety, with an emphasis on prevention of injury from falls and motor vehicle collisions.

Cara Hamann is a Clinical Assistant Professor for the University Of Iowa College Of Public Health, Department of Epidemiology. Her main research interest is transportation safety, with a focus on vulnerable and high-risk road users, including bicyclists, pedestrians, motorcyclists, and older and younger drivers.

Brandon Butcher is a graduate research assistant for the University of Iowa Injury Prevention Research Center and PhD student in the University Of Iowa College Of Public Health, Department of Biostatistics.

Corinne Peek-Asa is the associate dean for research for the University of Iowa College of Public Health and professor in the Department of Occupational and Environmental Health. Dr. Peek-Asa is the Director for the University of Iowa Injury Prevention Research Center.

Journal of Safety Research 76 (2021) 64–72

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

Modeling crash severity by considering risk indicators of driver and roadway: A Bayesian network approach

Yanchao Song, Siyuan Kou, Chen Wang *

Intelligent Transportation Systems Research Center, Southeast University, Nanjing 211189, China

ARTICLE INFO

Article history:
Received 20 January 2020
Received in revised form 22 June 2020
Accepted 18 November 2020
Available online 14 December 2020

Keywords:
Bayesian network
Information value
Crash severity
Risk indicators

ABSTRACT

Introduction: Traffic crashes could result in severe outcomes such as injuries and deaths. Thus, understanding factors associated with crash severity is of practical importance. Few studies have deeply examined how prior violation and crash experience of drivers and roadways are associated with crash severity. *Method:* In this study, a set of risk indicators of road users and roadways were developed based on their prior violation and crash records (e.g., cumulative crash frequency of a roadway), in order to reflect certain aspect or degree of their driving risk. To explore the impacts of those indicators on crash severity and complex interactions among all contributing factors, a Bayesian network approach was developed, based on citywide crash data collected in Kunshan, China from 2016 to 2018. A variable selection procedure based on Information Value (IV) was developed to identify significant variables, and the Bayesian network was employed to explicitly explore statistical associations between crash severity and significant variables. *Results*: In terms of balanced accuracy and AUCs, the proposed approach performed reasonably well. Bayesian modeling results indicated that the prior crash/violation experiences of road users and roadways were very important risk indicators. For example, migrant workers tend to have high injury risk due to their dangerous violation behaviors, such as retrograding, red-light running, and right-of-way violation. Furthermore, results showed that certain variable combinations had enhanced impacts on severity outcome than single variables. For example, when a migrant worker and a non-motorized vehicle are involved in a crash happening on a local road with high cumulative violation frequency in the previous year, the probability for drivers suffering serious injury or fatality is much higher than that caused by any single factor. *Practical applications:* The proposed methodology and modeling results provide insights for developing effective countermeasures to reduce crash severity and improve traffic system safety performance.

1. Introduction

The issue of traffic safety is a global problem that results in property damage, injury, and death, and costs billions of dollars every year. Statistical data show that approximately 50 million people were injured and about 1.35 million were killed in traffic crashes worldwide annually (Global Status Report on Roadway Safety, 2019). In China, there were about 203,000 traffic crashes occurred in 2017, resulting in about 64,000 deaths and approximately 175.8 million in direct property losses (China Statistical Yearbook, 2019). Road traffic crashes not only cause huge property losses and casualties of road users, but is also an important factor determining whether the road traffic system can perform well.

The road traffic system is a complex system composed of driver, vehicle, roadway, and other elements. The severity of traffic crashes will also be affected by various factors in the traffic system. When several risk factors exist simultaneously, there will be a higher risk of traffic crashes (Keller & Modarres, 2005; Xu, Wang, Liu, Wang, & Bao, 2018). Factors affecting severity of traffic crashes are usually caused by one or more of the following factors: driver characteristics, vehicle characteristics, roadway characteristics, crashes characteristics, and atmospheric factors (Chang & Wang, 2006; Coeliac, Papadimitriou, Papandreou, & Prevedouros, 2007; de Oña, Mujalli, & Calvo, 2011; Delen, Sharda, & Bessonov, 2006; Guo, Li, Wu, & Xu, 2018a, 2018b; Híjar, Carrillo, Flores, Anaya, & Lopez, 2000; Wang et al., 2017, 2018, 2019). However, very few studies have further explored how prior violation and crash experience of drivers and roadways are associated with crash severity, such as the cumulative violation frequency of a driver, the cumulative crash frequency of a roadway, and so forth. Those factors, reflecting certain aspects and degree of crash risk, could also have an impact on the severity of traffic crashes.

* Corresponding author.
E-mail address: wkobec@hotmail.com (C. Wang).

https://doi.org/10.1016/j.jsr.2020.11.006

Regarding modeling techniques, statistical models have been extensively developed for exploring contributing factors on crash severity. Regression models have been widely developed to determine the contributing factors that cause a specific crash severity. The logistic regression model and the ordered probit model are most commonly used in traffic crash analysis (Al-Ghamdi, 2002; Bedard, Guyatt, Stones, & Hirdes, 2002; Kockelman & Kweon, 2002; Milton, Shankar, & Mannering, 2008; Yamamoto & Shankar, 2004; Yau, Lo, & Fung, 2006). However, most of these models are established based on certain assumptions and pre-defined underlying relationships (such as linear relationships) between dependent and independent variables, which could lead to biased estimations when these assumptions are violated (Chang & Wang, 2006). Recently, Bayesian network (BN) method has been popularly applied in traffic crash modeling research (Borg, Bjelland, & Njå, 2014; Chen et al., 2015; Hänninen, 2014; Hossain & Muromachi, 2012; Mbakwe, Saka, Choi, & Lee, 2016; Zhao, Wang, & Qian, 2012). In a BN model, the model structure can be defined by the network topological representation and the quantitative relationships among variables can be specified by the conditional probabilities. BNs are capable of modeling inter-correlated independent variables to better interpret their heterogeneous influence on severity outcomes from the attribute changes in crashes.

In light of these, this study aims to explore the relationship between risk indicators (e.g., factors related to prior violation and crash records of driver and roadway) and crash severity, as well as examine complex interactions among all contributing factors. A set of risk indicators were created based on prior violation and crash records and BN model was used for crash modeling. This paper is organized as follows: the previous studies are review in Section 2; the data description and pre-processing are provided in Section 3; followed by the methodology in Section 4; the model results and discussions are presented in Section 5; and the research effort is concluded in Section 6.

2. Literature review

In previous studies, efforts toward crash severity analyses have been substantial. The main objective of those was to identify factors that significantly affect the severity of traffic crashes. The crash severity is usually influenced by one or more of the following factors: driver characteristics, vehicle characteristics, roadway characteristics, crashes characteristics and atmospheric factors. Several studies found that drivers of electric vehicles and bicycles are high-risk groups causing serious traffic crashes (Coeliac et al., 2007; Guo et al., 2018; Wang, Xu, Xia, Qian, & Lu, 2018). Vehicle types were found to be associated with crash severity, and pedestrians, motorcycle and bicycle riders were identified to have higher risks of being injured than others in traffic crashes (Chang & Wang, 2006). The use of a restraint system like seat belts, use of alcohol or drugs, drivers' age and gender, and vehicle role in the crash were found to have an important influence on the severity of crashes, and weather conditions or the time of the crash did not seem to affect the severity risk of injury (Delen et al., 2006). Road types, lighting conditions, and weather conditions are also important factors affecting the severity of crashes (Híjar et al., 2000). The combination of certain risk factors can cause more serious crashes, such as frontal collision, no lighting, and young drivers (de Oña et al., 2011; Wang, Xu, & Dai, 2019). Many previous studies have shown that the characteristics of driver and roadway in the traffic system are related to the severity of crashes. However, very few studies have further explored prior violation and crash records of driver and roadway, such as the cumulative violation frequency committed by the driver in the previous year, the cumulative crash frequency happened on the roadway in the previous year, and so forth. These significant indicators of the high-risk characteristics of driver and roadway could also have an impact on the severity of crashes.

Various methodological and statistical modeling techniques have been developed to explore contributing factors on crash severity. Regression analysis has been widely conducted to identify the contributing factors that cause a specific crash severity. The logistic regression model and the ordered probit model are most commonly applied in traffic crash analysis (Al-Ghamdi, 2002; Bedard et al., 2002; Kockelman & Kweon, 2002; Milton et al., 2008; Yamamoto & Shankar, 2004; Yau et al., 2006). Al-Ghamdi developed a logistic regression model to examine the contribution of several factors to crash severity (Al-Ghamdi, 2002). Milton et al. developed a mixed logit model for highway crash severity (Milton et al., 2008). Bedard et al. conducted a multivariate logistic regression to examine the independent contribution of driver, crash, and vehicle characteristics to fatal injuries (Bedard et al., 2002). Yau et al. applied a stepwise logistic regression model to identify risk factors for severe traffic crashes (Yau et al., 2006). Yamamoto and Shankar developed a bivariate ordered-response probit model of driver's and most severely injured passenger's severity in collisions with fixed objects (Yamamoto & Shankar, 2004). Kockelman and Kweon applied an ordered probit model to examine the risk of different injury levels sustained under all crash types, two-vehicle crashes, and single-vehicle crashes (Kockelman & Kweon, 2002). However, most of these models have certain assumptions on data structure/variable relationships. For instance, MNL models assume that random errors are Gumbel-distributed. Linear models assume a linear relationship between depended and independent variables. When being violated, such assumptions may lead to biased estimations (Chang & Wang, 2006).

Recently, Bayesian network (BN) method has been popularly applied in traffic crash modeling research (Borg et al., 2014; Chen et al., 2015; Hänninen, 2014; Hossain & Muromachi, 2012; Mbakwe et al., 2016; Zhao et al., 2012). Zhao et al. investigated the factors that affect hazardous material transportation crashes by developing a BN model based on related expert knowledge (Zhao et al., 2012). Borg et al. proved the applicability and effectiveness of a BN model on tunnel risk assessments (Borg et al., 2014). Chen et al. developed a hybrid approach to combine multinomial logit models and BN methods for comprehensively analyzing driver injury severity in crashes (Chen et al., 2015). Hossain and Muromachi applied Bayesian belief net to build the real-time crash prediction model (Hossain & Muromachi, 2012). Mbakwe et al. developed a hybrid approach to combine Delphi techniques and BN methods for modeling highway traffic crashes and forecasting crash rates in the countries of research (Mbakwe et al., 2016). Hänninen discussed the utilization of BNs in maritime safety modelling (Hänninen, 2014). The comprehensive understandings of BN applications in traffic crash modeling and analysis were provided in these studies. A BN model not only captures statistical associations between independent and dependent variables, but also extract correlations among independent variables to present all interactions among all variables, which can make it easy to describe crashes that involve many interdependent variables. Without any pre-defined assumptions on certain data strucuture (e.g., normal distribution) or variable relationships (e.g., a linear relationship), the structure of BN can be directly learned from crash data. With BN, the mechanism of traffic crashes can be visualized in the form of directed acyclic graph.

3. Data

3.1. Data description

This study was conducted based on three-year crash data and violation data collected in Kunshan, China from 2016 to 2018, obtained from the Kunshan Traffic Police Department. The crash data consist of crash records, crash characteristics, vehicle characteristics, roadway characteristics, and drivers' demographic information. Drivers' violation records were also collected by matching IDs in the crash data and violation data. This study focused on the impact of violation and crash records of driver and roadway in the previous year on the severity of crashes. Due to the failure to obtain prior records of driver and roadway for the crash occurred in 2016, two-year data from 2017 to 2018, including 81,336 crashes, were used as input data of the model. The variable, SEV, was defined to categorize severity of crashes into three levels: property damage only (PDO), slight injury (SI), serious injury or fatality (SIF).

3.2. Data preprocessing

In this study, some variables were created as risk indicators of drivers and roadways based on prior violation and crash records, such as specific violation types, cumulative crash frequency/violation frequency/violation penalty points/violation penalty fee of driver, cumulative crash frequency/crash casualties/crash property loss/violation frequency/violation penalty points/violation penalty fee of roadway. For processing continuous variables is computationally expensive due to their estimation inefficiency, these variables were discretized by a supervised discretization method (i.e., MODL method) which can potentially degrade structure learning and parameter estimation (Boullé, 2004). Based on the discrete results, this study defined high-risk variables as the risk indicators of driver and roadway. In the next section, a Bayesian network will be developed based on the hypothesis that the proposed risk indicators are related to crash severity. The risk indicators are shown in Table 1. The description of all variables in the model and their abbreviations are shown in Table 2.

In this study, datasets are unbalanced by the limited number of high-severity crashes, which could cause model overfitting. Thus, ADASYN, a widely used over-sampling method, was introduced to deal with this issue. ADASYN is based on the idea of adaptively generating minority data samples according to their distributions: more synthetic data is generated for minority class samples that are harder to learn compared to those minority samples that are easier to learn. The ADASYN method can not only reduce the learning bias introduced by the original imbalance data distribution, but can also adaptively shift the decision boundary to focus on those difficult to learn samples. The ADASYN algorithm is described below (He, Bai, Garcia, & Li, 2008).

Input

Table 1
The risk indicators.

Variable	Definition
High cumulative crash frequency of driver involvement	One crash or more
High cumulative violation frequency of driver involvement	One violation or more
High cumulative violation penalty points of driver involvement	Two scores or more
High cumulative violation penalty fee of driver involvement	More than 90 RMB
High cumulative crash frequency of roadway	More than 1300 crashes
High cumulative crash casualties of roadway	More than 120 people
High cumulative crash property loss of roadway	More than 5,000,000 RMB
High cumulative violation frequency of roadway	More than 3600 violations
High cumulative violation penalty points of roadway	More than 7800 scores
High cumulative violation penalty fee of roadway	More than 270,000 RMB

Table 2
Variable definitions and abbreviations.

Feature	Variable	Abbr
Gender	Female driver involvement	FDI
Age	Young driver (age < 18) involvement	YDI
	Middle-age driver (18–55) involvement	MDI
	Older driver (age > 55) involvement	ODI
Occupation	Migrant worker involvement	MWI
	Courier involvement	CRI
	Student involvement	SDI
Violation behavior	Retrograde violation of motorized vehicle involvement	RVMI
	Retrograde violation of non-motorized vehicle involvement	RVNI
	Right-of-way violation of motorized vehicles involvement	RWVMI
	Right-of-way violation of non-motorized vehicles involvement	RWVNI
	Traffic signal violation of motorized vehicles involvement	TSVMI
	Traffic signal violation of non-motorized vehicles involvement	TSVNI
	Drunk driving violation of motorized vehicles involvement	DDVMI
	Drunk driving violation of non-motorized vehicles involvement	DDVNI
	Speeding violation of motorized vehicles involvement	SVMI
	Speeding violation of non-motorized vehicles involvement	SVNI
	Overtaking violation of motorized vehicles involvement	OVMI
	Overtaking violation of non-motorized vehicles involvement	OVNI
Risk indicators of driver	High cumulative crash frequency of driver involvement	HCCFDI
	High cumulative violation frequency of driver involvement	HCVFDI
	High cumulative violation penalty points of driver involvement	HCVPPDI
	High cumulative violation penalty fee of driver involvement	HCVPFDI
Vehicle type	Non-motorized vehicles involvement	NMVI
	Motorcycle involvement	MCI
Roadway type	Urban expressway	UE
	Major arterial	MAA
	Minor arterial	MIA
	Local road	LR
	High-grade highway	HH
	Middle-grade highway	MH
	Low-grade highway	LH
Risk indicators of roadway	High cumulative crash frequency of roadway	HCCFR
	High cumulative crash casualties of roadway	HCCCR
	High cumulative crash property loss of roadway	HCCPLR
	High cumulative violation frequency of roadway	HCVFR
	High cumulative violation penalty points of roadway	HCVPPR
	High cumulative violation penalty fee of roadway	HCVPFR
Time period	Late night (00:00–05:00)	LN
	Morning peak hour (05:00–09:00)	MPH
	Off-peak hour (09:00–17:00)	OPH
	Evening peak hour (17:00–21:00)	EPH
	Night (21:00–00:00)	NT
Day	Weekdays	WD
Weather	Clear weather	CW
	Rainy weather	RW
	Overcast weather	OW
	Snowy weather	SW
	Fog weather	FW

Training data set D_{tr} with m samples $\{x_i, y_i\}$, $i = 1, \ldots, m$, where x_i is an instance in the n dimensional feature space X and $y_i \in Y = \{1, -1\}$ is the class identity label associated with x_i. Define m_s and m_l as the number of minority class examples and the number of majority class examples, respectively. Therefore, $m_s \leqslant m_l$ and $m_s + m_l = m$.

Procedure

(1) Calculate the degree of class imbalance:

$$d = m_s/m_l \tag{1}$$

where $d \in (0, 1]$.

(2) If $d < d_{th}$ then (d_{th} is a preset threshold for the maximum tolerated degree of class imbalance ratio):

(a) Calculate the number of synthetic data examples that need to be generated for the minority class:

$$G = (m_l - m_s) \times \beta \tag{2}$$

where $\beta \in [0, 1]$ is a parameter used to specify the desired balance level after generation of the synthetic data. $\beta = 1$ means a fully balanced data set is created after the generalization process.

(b) For each example $x_i \in minorityclass$, find K nearest neighbors based on the Euclidean distance in n dimensional space, and calculate the ratio r_i defined as:

$$r_i = \Delta_i/K, i = 1, \ldots, m_s \tag{3}$$

where Δ_i is the number of examples in the K nearest neighbors of x_i that belong to the majority class, therefore $r_i \in [0, 1]$.

(c) Normalize r_i according to $\hat{r}_i = r_i/\sum_{i=1}^{m_s} r_i$, so that $\hat{r}_i$ is a density distribution ($\sum_i \hat{r}_i = 1$).

(d) Calculate the number of synthetic data examples that need to be generated for each minority example x_i:

$$g_i = \hat{r}_i \times G \tag{4}$$

where G is the total number of synthetic data examples that need to be generated for the minority class as defined in Equation (2).

(e) For each minority class data example x_i, generate g_i synthetic data examples according to the following steps:

Do the **Loop** from 1 to g_i:

(i) Randomly choose one minority data example, x_{zi}, from the K nearest neighbors for data x_i.

(ii) Generate the synthetic data example:

$$s_i = x_i + (x_{zi} - x_i) \times \lambda \tag{5}$$

where $(x_{zi} - x_i)$ is the difference vector in n dimensional spaces, and λ is a random number: $\lambda \in [0, 1]$.

End **Loop**.

4. Methodology

4.1. Research design

BN is a technique for graphically representing a joint probability distribution of a selected set of variables (Pearl, 2014). The complicated relationships and interactions among independent and dependent variables can be visualized with graphical representations of BNs. The structure of a BN model is a directed graph, where the nodes represent the model variables and the links between the nodes represent the dependencies. Considering a large number of independent variables, BN structure optimization in the global space is extremely computation intensive. The search space increases as a super-exponential function of the number of variables (Chen et al., 2015). Therefore, a variable selection procedure is essential to extract a set of significant contributing variables and screen out variables that do not influence model performance, which also improves the interpretability of the model.

A variable selection procedure based on information value is applied to extract significant variables and screen out unnecessary and redundant variables to increase optimal BN structure search efficiency. The details of the variable selection procedure and BN modeling are provided in the following sections.

4.2. Variable selection based on information value

Most of the feature selection methods, such as Focus, Sch193 and MIFES1, cannot be applied to large dataset because of the memory and/or computational complexity they impose (Dash & Liu, 1997). Additionally, some of the methods for feature selection are applicable only on binary classification problems, such as Relief and Sege84 (Dash & Liu, 1997). In order to address these issues, the measure named information value (IV) could be used which can rank variables on the basis of their importance. Information value can be used for feature selection is less demanding in terms of memory and computational power. Owing to its convenient rules and fast calculation, IV is a popular and widely used measure in variable selection (Anderson, 2007; Finlay, 2012; Mays & Lynas, 2004). The IV tells the predictive power of an independent variable in relation to the dependent variable. Since it evolved from credit scoring world, it is generally described as a measure of the separation of good and bad customers. The formula for IV is shown below.

$$IV_{ji} = (\frac{B_{ji}}{B_j} - \frac{G_{ji}}{G_j}) \times \ln\frac{B_{ji}/B_j}{G_{ji}/G_j} \tag{6}$$

$$IV_j = \sum_i^n IV_{ji} \tag{7}$$

$$IV = \sum_j^m IV_j \times W_j \tag{8}$$

where

G_{ji}: Total goods of classification j of dependent variable in relation to classification i of independent variable.
B_{ji}: Total bads of classification j of dependent variable in relation to classification i of independent variable.
G_j: Total goods of classification j of dependent variable.
B_j: Total bads of classification j of dependent variable.
IV_{ji}: IV of classification$i$of independent variable in relation to classification j of dependent variable.
IV_j: IV of independent variable in relation to classification j of dependent variable.
W_j: The weight of classification j of dependent variable.
IV: IV of independent variable in relation to dependent variable.

There is reliable predictive power of an independent variable in relation to the dependent variable when IV is more than 0.1 (Siddiqi, 2012; Zdravevski, Lameski, Kulakov, & Gjorgjevikj, 2014). IVs were calculated of all contributing variables in Table 2 in relation to dependent variable, SEV. These significant variables with IV more than 0.1 will be extracted for BN structure establishment and probabilistic parameter learning to explicitly formulate statistical dependence between crash severity and explanatory attributes.

4.3. Bayesian network model

BN is employed as a classifier to analyze crash severity based on the significant factors identified in the variable selection procedure in this study. Let $U = \{x_1, x_2, \cdots, x_n\}$, $n \geqslant 1$ be a set of variables. A BN over a set of variables U is a network structure, which is a Directed Acyclic Graph (DAG) over U and a set of probability tables

$B_P = \{p(x_i|pa(x_i), x_i \in U)\}$ where $pa(x_i)$ is the set of parents or antecedents of $x_i$in BN and $i = 1, 2, 3, \cdots, n$. A BN represents joint probability distributions $P(U) = \prod_{x_i \in U} p(x_i|pa(x_i))$. The modeling process of BNs consists of structure learning and parameter learning.

There are two main approaches to structure learning in BNs: constraint-based algorithms (e.g., grow-shrink, incremental association markov blanket) and score-based algorithms (e.g., hill climbing, tabu search). With less sensitive to errors in individual tests, score-based algorithms have an advantage over the constraint-based algorithms. The hill climbing algorithm was applied in this study mainly because it is fast and widely used, and also produces good results in terms of network complexity and accuracy (Madden, 2008). The identified optimal BN structure presents the statistically dependent relationships among the variables in the model based on the dataset. The conditional probability tables can be estimated with Bayesian parameter estimation method. Through structure learning and parameter learning, BNs are developed to investigate the impacts of significant contributing variables extracted on the severity of crashes.

5. Results and discussion

5.1. IV calculation and significant variable identification

In this study, the significant variables were identified based on IV of all independent variables in relation to dependent variable, SEV. The significant variables with IV more than 0.1 were extracted as shown in Table 3.

As we can see in Table 3, 15 significant variables were identified from all variables. The IVs of MWI, NMVI are 0.395 and 0.487, respectively, indicating that the migrant worker and non-motorized vehicle involved in crashes are significant contributors to severity of crashes. The IV of HCCFR is 0.440, indicating that the roadway with high cumulative crash frequency is also a significant contributor to severity of crashes.

5.2. BN model comparison and discussion

The whole dataset was divided into two subsets, approximately 3:1. The larger subset (training dataset) was used for BN model learning and the smaller subset (testing dataset) was utilized for model validation and performance test.

This study primarily focused on the impact of risk indicators of drivers and roadways on crash severity. Therefore, two BN models, model A and model B, were developed in this study. Compared to model A, model B contains risk indicators of driver and roadway. BN structures of two models are shown in Fig. 1 and Fig. 2, respectively. Model A contains eleven variables and model B contains 16 variables. Node SEV has six parents (i.e., MWI, CRI, NMVI, MCI, LR, LH) in model A, and nine parents (i.e., MWI, NMVI, MCI, LR, LH, HCVPPDI, HCCFR, HCCCR, HCVFR) in model B. As can be seen in Fig. 2, the risk indicators of driver and roadway have direct impacts on the severity of crashes.

The concept of balanced accuracy was introduced to compute overall accuracy of model, which avoids inflated performance estimates on imbalanced datasets. It is the macro-average of recall scores per class or, equivalently, raw accuracy where each sample is weighted according to the inverse prevalence of its true class.

If y_i is the true value of the i-th sample, and w_i is the corresponding sample weight, then we adjust the sample weight to:

$$\hat{w}_i = \frac{w_i}{\sum_j 1(y_j = y_i) w_j} \tag{9}$$

Given predicted $\hat{y}_i$ for sample i, balanced accuracy is defined as:

$$balanced - accuracy(y, \hat{y}, w) = \frac{1}{\sum \hat{w}_i} \sum_i 1(\hat{y}_i = y_i)\hat{w}_i \tag{10}$$

Receiver Operating Characteristic (ROC) curve is another important indicator to evaluate the overall performance of the BN model. The overall ROC curves illustrate the tradeoffs between the true positive rates and false positive rates. The Area Under an ROC Curve (AUC) is a quantitative index that assesses the overall performance of model classification estimation. For multilabel classification, the assumption that all classes are equally important is often untrue. Therefore, the macro-averaging method was applied to calculate overall AUC of model (Tsoumakas & Vlahavas, 2007). In problems where infrequent classes, such as serious injury or fatality (SIF), are nonetheless important, macro-averaging may be a means of highlighting their performance. ROC curves of training and testing datasets of two models are shown in Fig. 3. The performance of two models are summarized in Table 4.

As can be seen in Table 4, the overall balanced accuracy of model A for training dataset and testing dataset are 65.4% and 64.8%, respectively. And the overall balanced accuracy of model B are 71.6% and 70.8%, respectively. The variance between the balanced accuracy for training dataset and testing dataset of two models are both under 1%, indicating that the trained networks are transferable and able to explain and model the testing data reasonably well. The variance between the AUCs for training dataset

Table 3
Significant variable identification.

Abbr	Significant variable	IV
ODI	Older driver (age > 55) involvement	0.144
MWI	Migrant worker involvement	0.395
CRI	Courier involvement	0.136
RVNI	Retrograde violation of non-motorized vehicle involvement	0.128
RWVNI	Right-of-way violation of non-motorized vehicles involvement	0.129
TSVNI	Traffic signal violation of non-motorized vehicles involvement	0.107
HCVPPDI	High cumulative violation penalty points of driver involvement	0.102
NMVI	Non-motorized vehicles involvement	0.487
MCI	Motorcycle involvement	0.140
LR	Local road	0.191
LH	Low-grade highway	0.238
HCCFR	High cumulative crash frequency of roadway	0.440
HCCCR	High cumulative crash casualties of roadway	0.117
HCVFR	High cumulative violation frequency of roadway	0.138
HCVPFR	High cumulative violation penalty fee of roadway	0.132

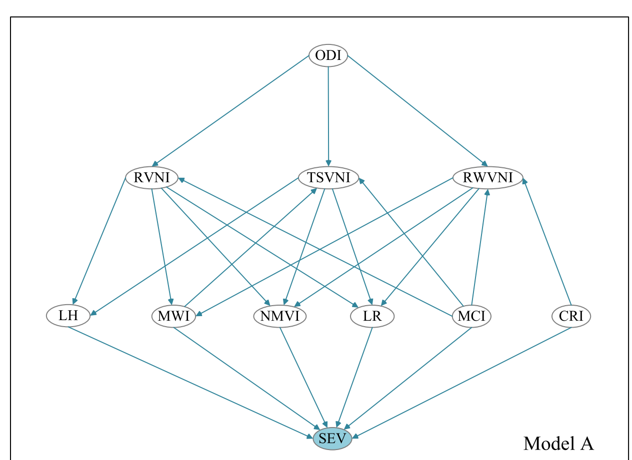


Fig. 1. BN structure of model A.

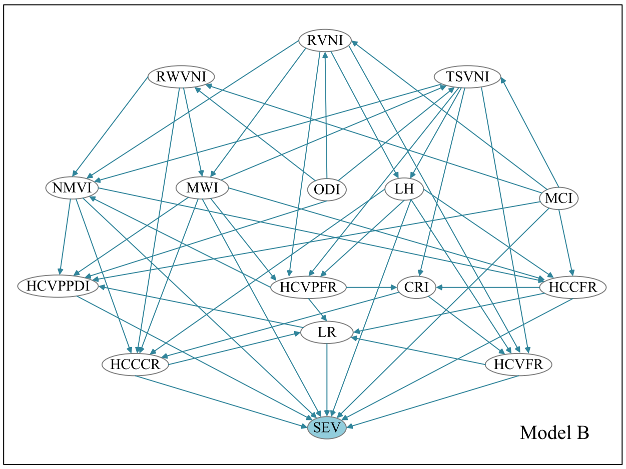


Fig. 2. BN structure of model B.

and testing dataset of two models are both under 0.02, indicating that two trained BNs are both not overfitting and perform well for three crash severity classifications. In terms of achieving a higher AUC, model B appeared to be better than model A, indicating that risk indicators of drivers and roadways have significant contribution to crash severity levels. In terms of both balanced accuracy and AUCs, model B with risk indicators included is a better model among all.

The BN structure explicitly formulates the interdependency among the variables and is capable of providing probability inference analyses based on the conditional probability tables for each node. Tables 5 and 6 illustrate probability inference results for nodes which have parents.

As can be seen in Fig. 2, node SEV have nine parents (i.e., MWI, NMVI, MCI, LR, LH, HCVPPDI, HCCFR, HCCCR, HCVFR) which have direct influences on crash severity. Through setting evidences for

Table 4
BN estimation results.

Model	Training dataset		Testing dataset	
	Balanced-accuracy	AUC	Balanced-accuracy	AUC
A	65.4%	0.820	64.8%	0.834
B	71.6%	0.860	70.8%	0.874

those parents, their contributions to crash severity can be quantified. Overall proportion distribution of three crash severities in the BN model are 0.325, 0.298, and 0.377, respectively. Table 6 illustrates probability inference results for variables which directly influencing crash severity.

As shown in Table 6, if a migrant worker (MWI) is involved in a crash, the probability for drivers suffering serious injury or fatality (SIF) increases from 0.377 to 0.409 (increased by 8.5%). This implies that migrant workers are more likely to be seriously injured in crashes. In traditional statistical framework, it could be difficult to further discuss why migrant workers could be injury prone. Thanks to the BN structure, node MWI can be seen to directly link to node HCVPPDI, RWVNI, TSVNI and RVNI. Thus, it was found that those road user group is associated with a large number of violation behaviors, especially red-light running, retrograde violation, and right-of-way violation. Moreover, those people are associated with roadways with high violation and crash frequency (HCVFR, HCCFR), indicating their high exposure. It was known that migrant workers normally lack educational background and safety awareness. Thus, those road user group needs to be paid more attention to.

As can be expected, if a non-motorized vehicle (NMVI) or a motorcycle (MCI) is involved in a crash, the likelihoods for drivers suffering serious injury or fatality (SIF) increase from 0.377 to 0.392 (increased by 4.0%) and from 0.377 to 0.394 (increased by 4.5%), respectively. This implies that drivers of non-motorized vehicles and motorcycles are more likely to be seriously and fatally injured in crashes, which is consistent with the conclusions in the

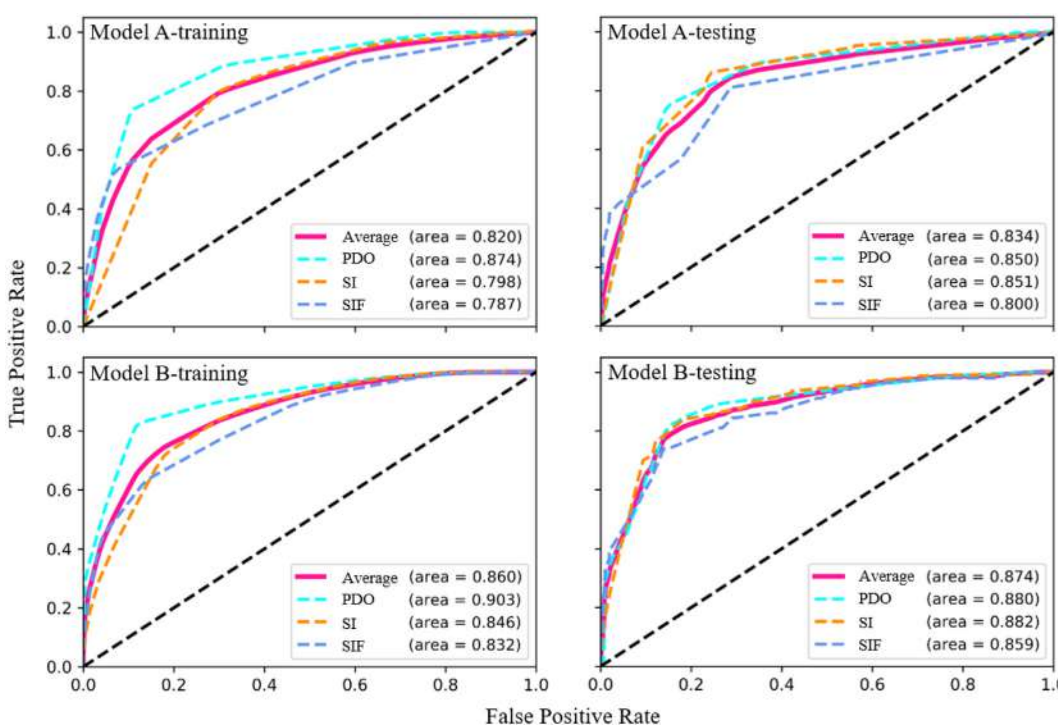


Fig. 3. ROC curves of training and testing datasets of two models.

Table 5
BN probability inference results for nodes.

Nodes	Parents	NO-NO	NO-YES	YES-NO	YES-YES
MWI	RVNI	0.858	0.908	0.142	0.092
	RWVNI	0.899	0.867	0.101	0.133
CRI	TSVNI	0.981	0.979	0.019	0.021
	HCCFR	0.962	0.998	0.038	0.002
	HCVPFR	0.975	0.984	0.025	0.016
RVNI	ODI	0.947	0.905	0.053	0.095
	MCI	0.867	0.985	0.133	0.015
RWVNI	ODI	0.929	0.847	0.071	0.153
	MCI	0.855	0.922	0.145	0.078
TSVNI	ODI	0.931	0.905	0.069	0.095
	MWI	0.940	0.896	0.060	0.104
	MCI	0.859	0.977	0.141	0.023
HCVPPDI	ODI	0.850	0.881	0.150	0.119
	MWI	0.872	0.859	0.128	0.141
	NMVI	0.838	0.893	0.162	0.107
	MCI	0.845	0.886	0.155	0.114
	LR	0.908	0.823	0.092	0.177
NMVI	RVNI	0.125	0.001	0.875	0.999
	RWVNI	0.126	0.000	0.874	1.000
	TSVNI	0.125	0.001	0.875	0.999
	HCVPFR	0.070	0.056	0.930	0.944
LR	HCCFR	0.546	0.565	0.454	0.435
	HCCCR	0.522	0.590	0.478	0.410
	HCVFR	0.657	0.455	0.343	0.545
	HCVPFR	0.667	0.445	0.333	0.555
LH	RVNI	0.663	0.681	0.337	0.319
	TSVNI	0.607	0.737	0.393	0.263
HCCFR	MWI	0.891	0.962	0.109	0.038
	NMVI	0.900	0.953	0.100	0.047
	MCI	0.885	0.968	0.115	0.032
	LH	0.957	0.896	0.043	0.104
HCCCR	MWI	0.871	0.810	0.129	0.190
	CRI	0.865	0.815	0.135	0.185
	RWVNI	0.932	0.748	0.068	0.252
	NMVI	0.792	0.888	0.208	0.112
	LH	0.806	0.875	0.194	0.125
HCVFR	CRI	0.622	0.831	0.378	0.169
	RVNI	0.758	0.695	0.242	0.305
	TSVNI	0.806	0.647	0.194	0.353
	LH	0.601	0.851	0.399	0.149
HCVPFR	MWI	0.615	0.762	0.385	0.238
	RVNI	0.718	0.659	0.282	0.341
	TSVNI	0.751	0.626	0.249	0.374
	LH	0.573	0.804	0.427	0.196

Table 6
BN probability inference results for variables influencing crash severities.

Parents	Severity		
	PDO 0.325	SI 0.298	SIF 0.377
MWI	0.299	0.292	0.409
	−0.026	−0.006	0.032
	−8.0%	−2.0%	8.5%
HCVPPDI	0.349	0.309	0.342
	0.024	0.011	−0.035
	7.4%	3.7%	−9.3%
NMVI	0.290	0.318	0.392
	−0.035	0.020	0.015
	−10.8%	6.7%	4.0%
MCI	0.308	0.298	0.394
	−0.017	0.000	0.017
	−5.2%	0.0%	4.5%
LR	0.309	0.319	0.372
	−0.016	0.021	−0.005
	−4.9%	7.0%	−1.3%
LH	0.352	0.315	0.333
	0.027	0.017	−0.044
	8.3%	5.7%	−11.7%
HCCFR	0.353	0.305	0.342
	0.028	0.007	−0.035
	8.6%	2.3%	−9.3%
HCCCR	0.313	0.322	0.365
	−0.012	0.024	−0.012
	−3.7%	8.1%	−3.2%
HCVFR	0.302	0.304	0.394
	−0.023	0.006	0.017
	−7.1%	2.0%	4.5%

previous studies (Chang & Wang, 2006; Coeliac et al., 2007; Guo et al., 2018; Wang et al., 2018). According to BN, drivers of non-motorized vehicles and motorcycles are also associated with a large number of violation behaviors (i.e., RWVNI, TSVNI and RVNI). Moreover, such group is also vulnerable due to lack of protection from their vehicles. To note, couriers (CRI) contribute a lot to red-light running violations (TSVNI). Couriers, especially for food delivery by motorcycles, have been drastically increased in recent years in China. However, such group often violates traffic rules and needs to be emphasized for governing.

Drivers driving on local roads (LR) or low-grade highways (LH) were found to be more involved in crashes that resulted in slight injury (SI). The probabilities for drivers suffering slight injury (SI) increase from 0.298 to 0.319 (increased by 7.0%) and from 0.298 to 0.315 (increased by 5.7%), respectively. But the probabilities for drivers suffering serious injury or fatality (SIF) decrease from 0.377 to 0.372 (decreased by 1.3%) and from 0.377 to 0.333 (decreased by 11.7%), respectively. As can be seen in Fig. 2, node LR is directly linked to node HCCFR, HCCCR, HCVFR, and HCVPFR. Nowadays, traffic enforcement systems have been extensively implemented on urban arterials in Kunshan, China. However, enforcement on local roads is still lacking, causing more violations and crashes. Node LH is linked to node RVNI and TSVNI, indicating that retrograding and red-light running violations largely happen on low-grade highways. Such behaviors could increase injury risk. Due to the relatively low operating speed of low-grade highways, the crashes happening on such roads are more likely to cause property damage and slight injury, but not serious injury or fatality.

There are several interesting findings according to the results. Risk indicators of drivers and roadways significantly affect the probabilities of driver injury and fatality in crashes, supported by the conditional probability tables for these nodes. As can be seen in Table 6, if drivers with high cumulative violation penalty points (HCVPPDI) are involved in crashes, the probability of suffering serious injury or fatality (SIF) decreases from 0.377 to 0.342 (decreased by 9.3%). Also note that node HCVPPDI is not linked to those dangerous violation types, including node RWVNI, TSVNI, and RVNI. Thus, it could be concluded that drivers with prior dangerous violation records (i.e., red-light running, retrograding, and right-of-way violation) could be more dangerous than those who only commit more common violations.

It should be also noted that if crashes happen on roadways with prior high crash risk (i.e., HCCFR, HCCCR), the probabilities of drivers suffering serious injury or fatality (SIF) tend to decrease from 0.377 to 0.342 (decreased by 9.3%) and from 0.377 to 0.365 (decreased by 3.2%), respectively. This could be attributed to the safety improvement actions conducted by the Kunshan Traffic Police Department in recent years. Another possible reason could be that drivers tend to be more cautious when driving on such high-crash-risk roadways. On the other hand, if crashes happen on roadways with prior high cumulative violation frequency (HCVFR), the probability of serious injury or fatality (SIF) increases from 0.377 to 0.394 (increased by 4.5%). This implies that these roadways with rampant dangerous violation behaviors would increase the propensity of driver injury and fatality.

On the basis of the above analysis, we know that some variables have an impact on the severity of traffic crashes. Considering the impact of variable combinations to crash severity, three high-severity crash scenarios were discussed, as shown in Table 7. In scenario 1, when a migrant worker (MWI) and a non-motorized

Table 7
High-severity crash scenarios.

Scenario	Driver	Vehicle	Roadway	Risk indicator	Severity		
					PDO 0.325	SI 0.298	SIF 0.377
1	MWI	NMVI	LR	HCVFR	0.261	0.286	0.453
2	MWI	NMVI	LR	HCCCR	0.271	0.280	0.449
3	MWI	MCI	LR	HCVFR	0.292	0.293	0.415

vehicle (NMVI) are involved in a crash happening on a local road (LR) with high cumulative violation frequency in the previous year (HCVFR), the probability for drivers suffering serious injury or fatality (SIF) is as high as 0.453 (increased by 20.2%). In scenario 2, when a migrant worker (MWI) and a non-motorized vehicle (NMVI) are involved in a crash happening on a local road (LR) with high cumulative crash casualties in the previous year (HCCCR), the probability for drivers suffering serious injury or fatality (SIF) is as high as 0.449 (increased by 19.1%). In scenario 3, when a migrant worker (MWI) and a motorcycle (MCI) are involved in a crash happening on a local road (LR) with high cumulative violation frequency in the previous year (HCVFR), the probability for drivers suffering serious injury or fatality (SIF) is as high as 0.415 (increased by 10.1%). These scenarios indicate that some combinations of variables could have much larger impacts on severity outcome than any single variable.

6. Conclusions

In this study, some risk indicators of drivers and roadways were developed based on prior violation and crash records, which were believed to reflect certain aspects and degree of crash risk. A modeling approach was proposed to integrate a variable selection procedure based on information value and Bayesian network to quantitatively and graphically analyze crash severity patterns. Variable selection procedure based on information value was able to improve the structure of BN and reducing computation intensity, by removing irrelevant factors and retaining important factors. In terms of balanced accuracy and AUCs, the proposed BN modeling approach performed reasonably well for three crash severity levels. The following major conclusions can be drawn:

(1) Migrant workers, couriers, non-motorized vehicles, motorcycles, local roads, and low-grade highway can significantly influence crash severity. Those findings could be especially useful for traffic management in developing countries.
(2) According to the modeling results, risk indicators (e.g., dangerous violation behaviors, high cumulative crash frequency of roadway) were found to have significantly direct and indirect effects on the severity of traffic crashes. Thanks to BN, those indicators could reveal some hidden risk aspects of drivers and roadways. For example, migrant workers tend to have high injury risk due to dangerous violation behaviors, such as retrograding, red-light running, and right-of-way violation.
(3) Some combinations of variables have enhanced impacts on severity outcome compared to single variables. For example, when a migrant worker and a non-motorized vehicle are involved in a crash happening on a local road with high cumulative violation frequency in the previous year, the probability for drivers suffering serious injury or fatality (SIF) is much higher than that caused by any single factor.

In general, the proposed approach appears to be a promising and reliable tool to identify factors significantly increasing crash severity, and the study findings provide insights for developing effective countermeasures to reduce severe crashes and improve traffic system safety performance. Some issue need to be further addressed. For example, since only IV was used for feature selection in this study, other methods (e.g., random forest) can also be investigated in a future study.

Acknowledgements

This research has been supported by the National Key R&D Program of China (2018YFE0102700) and National Natural Science Foundation of China (71971061). The authors gratefully acknowledge assistance with crash data and violation data from Kunshan Traffic Police Department.

References

Al-Ghamdi, A. S. (2002). Using logistic regression to estimate the influence of accident factors on accident severity. *Accident Analysis & Prevention, 34*(6), 729–741.
Anderson, R. (2007). *The credit scoring toolkit: Theory and practice for retail credit risk management and decision automation*. Oxford University Press.
Bedard, M., Guyatt, G. H., Stones, M. J., & Hirdes, J. P. (2002). The independent contribution of driver, crash, and vehicle characteristics to driver fatalities. *Accident Analysis & Prevention, 34*(6), 717–727.
Borg, A., Bjelland, H., & Njå, O. (2014). Reflections on Bayesian Network models for roadway tunnel safety design: A case study from Norway. *Tunnelling and Underground Space Technology, 43*, 300–314.
Boullé, M. (2004). A Bayesian approach for supervised discretization. *WIT Transactions on Information and Communication Technologies, 33*.
Chang, L. Y., & Wang, H. W. (2006). Analysis of traffic injury severity: An application of non-parametric classification tree techniques. *Accident Analysis & Prevention, 38*(5), 1019–1027.
Chen, C., Zhang, G., Tarefder, R., Ma, J., Wei, H., & Guan, H. (2015). A multinomial logit model-Bayesian network hybrid approach for driver injury severity analyses in rear-end crashes. *Accident Analysis & Prevention, 80*, 76–88.
China Statistical Yearbook. National Bureau of Statistics. http://www.stats.gov.cn/tjsj/ndsj/2018/indexch.htm. Accessed August 1, 2019.
Coeliac, P., Papadimitriou, F., Papandreou, K., & Prevedouros, P. (2007). Urban freeway crash analysis: Geometric, operational, and weather effects on crash number and severity. *Transportation Research Record, 2015*(1), 123–131.
Dash, M., & Liu, H. (1997). Feature selection for classification. *Intelligent Data Analysis, 1*(3), 131–156.
de Oña, J., Mujalli, R. O., & Calvo, F. J. (2011). Analysis of traffic accident injury severity on Spanish rural highways using Bayesian networks. *Accident Analysis & Prevention, 43*(1), 402–411.
Delen, D., Sharda, R., & Bessonov, M. (2006). Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. *Accident Analysis & Prevention, 38*(3), 434–444.
Finlay, S. (2012). *Credit scoring, response modeling, and insurance rating: A practical guide to forecasting consumer behavior*. Palgrave Macmillan.
Global Status Report on Roadway Safety. World Health Organization. https://www.who.int/gho/road_safety/en/. Accessed August 1, 2019.
Guo, Y., Li, Z., Wu, Y., & Xu, C. (2018). Exploring unobserved heterogeneity in bicyclists' red-light running behaviors at different crossing facilities. *Accident Analysis & Prevention, 115*, 118–127.
Guo, Y., Li, Z., Wu, Y., & Xu, C. (2018). Evaluating factors affecting electric bike users' registration of license plate in China using Bayesian approach. *Transportation Research Part F: Traffic Psychology and Behaviour, 59*, 212–221.
Hänninen, M. (2014). Bayesian networks for maritime traffic accident prevention: Benefits and challenges. *Accident Analysis & Prevention, 73*, 305–312.
He, H., Bai, Y., Garcia, E. A., & Li, S. (2008). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)* (pp. 1322–1328). IEEE.
Híjar, M., Carrillo, C., Flores, M., Anaya, R., & Lopez, V. (2000). Risk factors in highway traffic accidents: A case control study. *Accident Analysis & Prevention, 32*(5), 703–709.
Hossain, M., & Muromachi, Y. (2012). A Bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways. *Accident Analysis & Prevention, 45*, 373–381.
Keller, W., & Modarres, M. (2005). A historical overview of probabilistic risk assessment development and its use in the nuclear power industry: A tribute to the late Professor Norman Carl Rasmussen. *Reliability Engineering & System Safety, 89*(3), 271–285.
Kockelman, K. M., & Kweon, Y. J. (2002). Driver injury severity: An application of ordered probit models. *Accident Analysis & Prevention, 34*(3), 313–321.
Madden, M. G. (2008). On the classification performance of TAN and general Bayesian networks. In *International conference on innovative techniques and applications of artificial intelligence* (pp. 3–16). London: Springer.
Mays, F. E., & Lynas, N. (2004). *Credit scoring for risk managers: The handbook for lenders*. Ohio: Thomson/South-Western.

Mbakwe, A. C., Saka, A. A., Choi, K., & Lee, Y. J. (2016). Alternative method of highway traffic safety analysis for developing countries using delphi technique and Bayesian network. *Accident Analysis & Prevention, 93*, 135–146.
Milton, J. C., Shankar, V. N., & Mannering, F. L. (2008). Highway accident severities and the mixed logit model: An exploratory empirical analysis. *Accident Analysis & Prevention, 40*(1), 260–266.
Pearl, J. (2014). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Elsevier.
Siddiqi, N. (2012). *Credit risk scorecards: developing and implementing intelligent credit scoring (Vol. 3)*. John Wiley & Sons.
Tsoumakas, G., & Vlahavas, I. (2007). Random k-labelsets: An ensemble method for multilabel classification. In *European conference on machine learning* (pp. 406–417). Berlin, Heidelberg: Springer.
Wang, C., Xu, C., & Dai, Y. (2019). A crash prediction method based on bivariate extreme value theory and video-based vehicle trajectory data. *Accident Analysis & Prevention, 123*, 365–373.
Wang, C., Xu, C., Xia, J., & Qian, Z. (2017). Modeling faults among e-bike-related fatal crashes in China. *Traffic Injury Prevention, 18*(2), 175–181.
Wang, C., Xu, C., Xia, J., Qian, Z., & Lu, L. (2018). A combined use of microscopic traffic simulation and extreme value methods for traffic safety evaluation. *Transportation Research Part C: Emerging Technologies, 90*, 281–291.
Xu, C., Wang, Y., Liu, P., Wang, W., & Bao, J. (2018). Quantitative risk assessment of freeway crash casualty using high-resolution traffic data. *Reliability Engineering & System Safety, 169*, 299–311.
Yamamoto, T., & Shankar, V. N. (2004). Bivariate ordered-response probit model of driver's and passenger's injury severities in collisions with fixed objects. *Accident Analysis & Prevention, 36*(5), 869–876.
Yau, K. K., Lo, H. P., & Fung, S. H. (2006). Multiple-vehicle traffic accidents in Hong Kong. *Accident Analysis & Prevention, 38*(6), 1157–1161.
Zdravevski, E., Lameski, P., Kulakov, A., & Gjorgjevikj, D. (2014). Feature selection and allocation to diverse subsets for multi-label learning problems with large datasets. In *2014 federated conference on computer science and information systems* (pp. 387–394). IEEE.
Zhao, L., Wang, X., & Qian, Y. (2012). Analysis of factors that influence hazardous material transportation accidents based on Bayesian networks: A case study in China. *Safety Science, 50*(4), 1049–1055.

Yanchao Song. **Position:** Graduate student in Intelligent Transportation Systems Research Center, Southeast University. **Education:** Master of Engineering Candidate in Transportation Engineering, Southeast University (2017–Present). Bachelor of Engineering in Traffic Engineering, Chang'an University (2013–2017). **Reaearch interest:** Traffic safety, driving behavior and traffic accident modeling.

Siyuan Kou. **Position:** Graduate student in Intelligent Transportation Systems Research Center, Southeast University. **Education:** Master of Engineering Candidate in Transportation Engineering, Southeast University (2018–Present). Bachelor of Engineering in Traffic Engineering, Chang'an University (2014–2018). **Reaearch interest:** Traffic safety, traffic data mining.

Chen Wang. **Position:** Professor in Intelligent Transportation Systems Research Center, Southeast University. **Education:** Ph.D. in Civil and Environmental Engineering, University of Kentucky (2009–2012) Master of Engineering in Transportation Engineering, Southeast University (2007–2010). Bachelor of Engineering in Transportation Engineering, Southeast University (2003–2007). **Reaearch interest:** Traffic safety, driving behavior, traffic accident modeling, traffic data mining.