Public Health 215 (2023) 42–55

Contents lists available at ScienceDirect

Public Health

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

Original Research

Ambient air pollution exposure and telomere length: a systematic review and meta-analysis

Z.-Q. Zong [a, †], S.-W. Chen [b, †], Y. Wu [c, d], S.-Y. Gui [a], X.-J. Zhang [e, **], C.-Y. Hu [e, f, *]

[a] Department of Clinical Medicine, The Second School of Clinical Medicine, Anhui Medical University, 81 Meishan Road, Hefei 230032, China
[b] Department of General Surgery, The First Affiliated Hospital of Anhui Medical University, 218 Jixi Road, Hefei 230022, China
[c] Oncology Department of Integrated Traditional and Western Medicine, The First Affiliated Hospital of Anhui Medical University, 218 Jixi Road, Hefei 230022, China
[d] The Integrated Traditional and Western Medicine Cancer Center of Anhui Medical University, 81 Meishan Road, Hefei 230032, China
[e] Department of Epidemiology and Biostatistics, School of Public Health, Anhui Medical University, 81 Meishan Road, Hefei 230032, China
[f] Department of Humanistic Medicine, School of Humanistic Medicine, Anhui Medical University, 81 Meishan Road, Hefei 230032, China

ARTICLE INFO

Article history:
Received 22 February 2022
Received in revised form
20 September 2022
Accepted 28 November 2022
Available online 13 January 2023

Keywords:
Air pollution
Telomere length
Children
Systematic review
Meta-analysis

ABSTRACT

Objective: This study aimed to provide evidence of the associations between pre- and post-birth and adulthood air pollution exposure with telomere length.
Study design: The databases of PubMed, Embase, and Web of Science were searched up to June 1st, 2022 in order to include relevant observational studies and perform a systematic review and meta-analysis.
Methods: The random-effects meta-analysis was grouped by air pollutant and exposure window (pre- and post-birth and adulthood) to evaluate the summary effect estimate. Cochran's Q and I^2 statistics were used to evaluate the heterogeneity among the included studies. The quality of individual studies was evaluated using the national toxicology program/office of health assessment and translation risk of bias rating tool.
Results: We identified 18 studies, covering 8506 children and 2263 adults from multiple countries. We found moderate evidence that particulate matter less than 2.5 μm ($PM_{2.5}$) exposure during the entire pregnancy (−0.043, 95% CI: −0.067, −0.018), nitrogen dioxide (NO_2) exposure during the first trimester (−0.016, 95% confidence interval [CI]: −0.027, −0.005), long-term adulthood $PM_{2.5}$ exposure were associated with shortening telomere length. Mild to high between-study heterogeneity was observed for the most tested air pollutant-telomere length combinations in different exposure windows.
Conclusions: This systematic review and meta-analysis provides the evidence which strongly supports that prenatal $PM_{2.5}$ and NO_2 exposures were related to reduced telomere length, while prenatal sulfur dioxide (SO_2) and carbon monoxide (CO) exposures, childhood $PM_{2.5}$, particulate matter less than 10 μm (PM_{10}), NO_2 exposures and short-term adulthood $PM_{2.5}$ and PM_{10} exposures were not associated with telomere length. Further high-quality studies are needed to elaborate our suggestive associations.

Introduction

Environmental pollution caused by ambient air has become one of the biggest global environmental issues. According to the recent Global burden of diseases report, air pollution accounts for more than one in nine of all deaths worldwide, contributing to 6.67 million deaths (95% confidence interval [CI]: 5.90 to 7.49 million) in 2019. Globally, air pollution was the fourth leading cause of death in 2019.[1] The major health problems associated with air pollution include ischemic heart disease, lung cancer, chronic obstructive pulmonary disease, lower respiratory infections (such as pneumonia), stroke, type two diabetes, and neonatal diseases primarily related to low birth weight and preterm birth.[2–10] The mechanisms by which air pollutants can negatively affect human health are hypothesized to be oxidative stress and inflammation.[11]

* Corresponding author. Department of Epidemiology and Biostatistics Anhui Medical University 81 Meishan Road, Hefei 230032, China.
** Corresponding author.
E-mail addresses: zhangxiujun@ahmu.edu.cn (X.-J. Zhang), cy.hu@ahmu.edu.cn (C.-Y. Hu).
† Zhi-Qiang Zong and San-Wei Chen contributed equally to this study.

https://doi.org/10.1016/j.puhe.2022.11.022
0033-3506/

There is evidence that environmental influences on telomere length are one intermediate step in linking air pollution with adverse health outcomes.[12,13] Cellular end caps known as telomeres are repeating, non-coding DNA-protein complexes that act as protective caps and contribute to genomic stability and chromosomal integrity.[14–16] In addition to shortening with every cell division, telomeres are highly susceptible to oxidative stress as a result of their guanine-enriched structure.[17] There has been an increasing body of evidence indicating that a range of pro-oxidant environmental exposures, such as ambient air pollution, may have an adverse impact on telomere attrition. It has been shown, for example, that smoking, black carbon, traffic-related air pollution, and particulate matter ($PM_{2.5}$) exposures in adults are associated with shorter telomeres.[12,13,18,19] Furthermore, some studies have examined this effect in children for the following reasons: (1) telomere length is a predictor of longevity, as demonstrated in the research of zebra finches,[20] (2) telomere length decreases rapidly during childhood,[21] and (3) telomere length varies significantly between individuals during childhood.[22]

There has been considerable research on the relationship between ambient air pollution exposure and telomere length over the past few years, but the evidence is inconclusive. For example, Wong et al. reported that cumulative $PM_{2.5}$ exposure was associated with a significant decrease of telomere length in workers exposure to welding fumes in the Boilermakers Study.[23] Additionally, Pieters et al. noted that each 5-$\mu g/m^3$ increase in annual $PM_{2.5}$ concentration was associated with a significant decrease in the length of the telomeres.[24] In contrast, Xia et al. found that short-term exposure to PM or gaseous pollutants did not appear to be associated with telomere length in patients with type two diabetes.[25] Several factors have contributed to the difficulty of comparing these results: most studies involved relatively few participants, and some studies included individuals who had experienced occupational exposures, and PM compositions in occupational settings may differ from those in environmental settings.

In addition, inconsistent results have also been reported for the association of exposure to ambient air pollution before and after birth with telomere length in offspring. Particularly, Walton et al. revealed that the telomere base pair content increased with the increased levels of NO_2, NO_x, $PM_{2.5}$, and PM_{10} exposure among school children,[26] whereas Moslem et al. reported that children exposed to specific solid substances in residences and kindergartens had shorter telomere lengths at higher exposure levels.[27] In a multicenter birth cohort study conducted in the European region, Clemente et al. found no significant association between prenatal $PM_{2.5}$ exposure and telomere length,[28] whereas Martens et al. reported that pregnant women exposed to higher concentrations of $PM_{2.5}$ delivered babies with shorter telomere length.[18] Furthermore, Rosa et al. identified gender and timing differences, namely, higher $PM_{2.5}$ levels during a particular fetal window were associated with shorter telomere length in girls than in boys.[29]

For a better understanding of the air pollution–telomere length associations, a comprehensive systematic review and meta-analysis is necessary in order to shorten the research gap. Nevertheless, to our knowledge, only one meta-analysis has been conducted on air pollution and telomere length in adults, which included a small number of studies.[30] In consideration of the fact that some new relevant studies have been published and the evidence regarding such associations on prenatal air pollution exposure and telomere length in offspring was not summarized. Thus, we conducted a systematic review and meta-analysis of human epidemiological studies to examine the association between ambient air pollution exposure and telomere length at both the pre- and post-birth stages of life, as well as during adulthood.

Methods

Study question

The search question was 'What is the impact of higher ambient air pollution exposure on telomere length compared to lower levels of air pollution exposure in the general population?'

Search strategy

In order to identify eligible studies from inception to June 1st, 2022, authors searched the Web of Science, PubMed, and Embase using the keywords that were representative of the exposure and outcomes listed in our PECOS (Population, Exposure, Comparator, Outcome, and Study design) statement ('air pollution' OR 'PM_{10}' OR '$PM_{2.5}$' OR 'CO' OR 'SO_2' OR 'NO_2' OR 'NO_x' OR 'air pollutant' OR 'particulate matter' OR 'sulfur dioxide' OR 'nitrogen dioxide' OR 'nitroxide' OR 'carbon dioxide') and ('telomere length' OR 'telomere' OR 'telomerase'). The outcomes of this study were structured and presented as claimed by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (Supplementary Material, Table S1).[31] The database search was restricted to original studies published in English. Additionally, references and related reviews of the identified articles were manually scanned, and a follow-up search was conducted prior to manuscript submission to identify qualified published data.

Study selection

According to a PECOS statement, the eligibility criteria for Population, Exposure, Comparator, Outcomes, and Study Design are as follows:

- Population—General population including children and adults.
- Exposure—Long-term and short-term exposure to ambient air pollution. The exposure time more than 30 days was considered as long-term exposure;[32] otherwise, short-term was considered. Exposure was expressed in continuous.
- Comparator—Reference categories identified as groups with lower exposure levels.
- Outcome—Telomere length change (telomere/single gene copy ratio).
- Study design—Cohort, case–control, and cross-sectional studies examining the association of air pollution with telomere length, which reported quantitative β coefficients and 95% CIs.

Accordingly, studies using only air pollution alternative indicators are excluded, such as distance from major roadways and traffic density of the nearest road. When studies had overlapping populations and information, we retained publication articles reporting the most comprehensive information or the most representative population.

Data extraction

Based on a pre-designed template, the following important information was extracted from eligible studies: (1) authors, year (country); (2) study design, population and statistical methods; (3) exposure assessment; (4) adjusted confounders; (5) main results. If one publication reports more than one β coefficient for the same outcome of interest, we extracted only the most adjusted effect estimates listed in included articles. The data extraction work was performed independently by two authors (ZQZ and SWC) and any disagreements were resolved by a group discussion.

Risk of bias assessment of individual study

The risk of bias was assessed for the studies included in the meta-analysis using the national toxicology program/office of health assessment and translation (NTP/OHAT) risk of bias rating tool for human and animal studies, which evaluates each study according to seven questions: (i) confounding bias, (ii) attrition/exclusion bias, (iii) detection bias for exposure, (iv) detection bias for outcome, (v) selective reporting bias, (vi) selective bias, and (vii) conflict of interest. The authors chose definitely low, probably low, probably high, or definitely high risk of bias for each question. Following that, the overall study quality was categorized into three tiers of categories (Tier 1, two, or 3). There is a detailed response instructions available on the official website (https://ntp.niehs.nih.gov/ntp/ohat/pubs/riskofbiastool_508.pdf).

Confidence in the body of evidence

Based on the GRADE approach,[33,34] the evidence quality for results in study was rated using the NTP/OHAT framework, 2019.[35] The aim of this study was to evaluate the evidence quality of research and the strength of recommendations based on Bradford–Hill criteria. Four descriptors ('very low', 'low' 'moderate', and 'high') were used to indicate the confidence level in the body of evidence. In the case of a 'high degree of confidence' conclusion, further investigation is unlikely to lead to confidence changes in the apparent relation between the substance exposure and the effect outcome. In order to determine an initial rating, it is necessary to determine whether exposure precedes, and is significantly associated with, the outcome of the study design.

The initial rating may be downgraded in the event if factors reduce the outcome confidence (unexplained inconsistency, risk of bias, indirectness, imprecision, and publication bias) or upgraded in the event if factors increase the outcome confidence (large magnitude of effect, dose response, consistency across study designs/populations/animal models, and consideration of residual confounding factors that increase confidence in the association).

Statistical analysis

We separated studies into those of pre- and post-birth air pollution exposure on telomere length in offspring, as well as air pollution exposure on telomere length of adults. To pool the effect estimates, we grouped each air pollutant ($PM_{2.5}$, NO_2, etc.) and telomere length together. If more than two studies were identified for the same combination, we used random-effects meta-analysis. As an indicator of such associations, the β coefficient was used in this study. We extracted β coefficients and 95% CIs from each included studies, with the most confounders adjusted model.

Among the studies that were considered meta-analyzable, exposures were transformed and scaled in different ways. Consequently, we re-scaled all exposure-outcome effect estimates to reflect a change in outcome per 10 $\mu g/m^3$ increase in $PM_{2.5}$, PM_{10}, and NO_2, 1 ppb in SO_2, and 100 $\mu g/m^3$ in CO for pre- and post-birth exposure, while for adulthood exposure, we re-scaled the effect estimates to each 5 $\mu g/m^3$ increase in $PM_{2.5}$ and PM_{10}. Based on a conversion factor of 0.75, we converted the NO_x effects into NO_2 effects.

If multiple effect estimates were reported, we selected the most representative effect estimate. In specific, Martens et al. measured the telomere length from cord blood and placental tissue at each exposure window,[18] we utilized the extracted β coefficients for the cord blood sample as the main analysis. In a study conducted by Moslem et al., they assessed childhood air pollution exposure using data collected at home, kindergarten-indoor, and kindergarten-outdoor, we used the data for the outcomes from residence/home air pollution exposure assessment.[27] In order to assess statistical heterogeneity between studies, we calculated Cochran's Q statistic ($P < 0.05$ for statistical significance) and I^2 values. In general, I^2 values range from 0 to 100% and are considered low if up to 25%, moderate if 25–50%, and high if above 50%.[36] Sensitivity analysis was performed using the leave-one-out method. The funnel plot in combination with Egger's test were used to assess the risk of publication bias. Stata 15.0 software was used to perform all the analyses in our study.

Results

Literature search and characteristics of included studies

As a result of searching the PubMed, Web of Science, and Embase library databases, 3661 results were found after 634 duplicates were removed. A preliminary screening of titles and abstracts identified 29 studies relating to air pollutants and the length of telomeres in the general population that were considered for further review. After reading the full-text of all potential eligible studies, 11 studies were excluded because they did not meet the inclusion criteria (Fig. 1). Finally, the present meta-analysis included 18 studies, out of which 12 focused on pre- and post-birth air pollution exposure,[18,26–29,37–43] and six focused on adulthood air pollution exposure.[23–25,44–46] The flow chart of study selection is shown in Fig. 1.

Summary characteristics of the 18 studies were shown in Table 1. Overall, studies were published between 2011 and 2022, with most studies (N = 10) were performed in Europe. For the studies on adults, the study populations were also varied, which included truck drivers,[46] steel workers,[45] police traffic officers,[13] and boiler makers.[23] In addition, the exposure windows also varied among the included studies. All the included studies used a relative unit (telomere/single gene copy ratio) for telomere length measurement. 16 studies investigated changes in telomere length using blood samples[18,23–25,27–29,37,39–46] and two studies using oral cells.[26,38] Quantitative polymerase chain reaction (qPCR) method was performed in three studies,[18,29,37] quantitative real-time polymerase chain reaction (qRT-PCR) method was used in 13 studies,[24,25,27,28,38,40–47] only two studies used monochrome multiplex quantitative polymerase chain reaction (MMq-PCR) method.[26,39]

Risk of bias assessment of the included studies

A total of 18 studies were rated as 'probably low risk of confounding bias'; one study was rated as 'probably high risk of confounding bias' because of the limited adjustment for confounding factors. Almost all studies (N = 17) were classified as 'probably low risk of detection bias' because exposure assessments were performed based on widely used models, except Carugno et al. used the monitoring station data.[44] All studies measured telomere length used blood sample, except Walton et al. and Hautekiet et al., which used an oral sample.[26,38] In addition, all studies did not report any evidence that outcome data were missing. In terms of selection bias, six studies on adults were classified as being 'probably high risk of bias,' as they recruited populations from occupational or highly exposed groups.[23–25,44–46] The included studies were funded by public funds and no interest conflict was reported by any of the authors. Accordingly, all studies were classified as either Tier 1 (N = 16) or Tier 2 (N = 2), which indicates that the risk of bias is probably low.

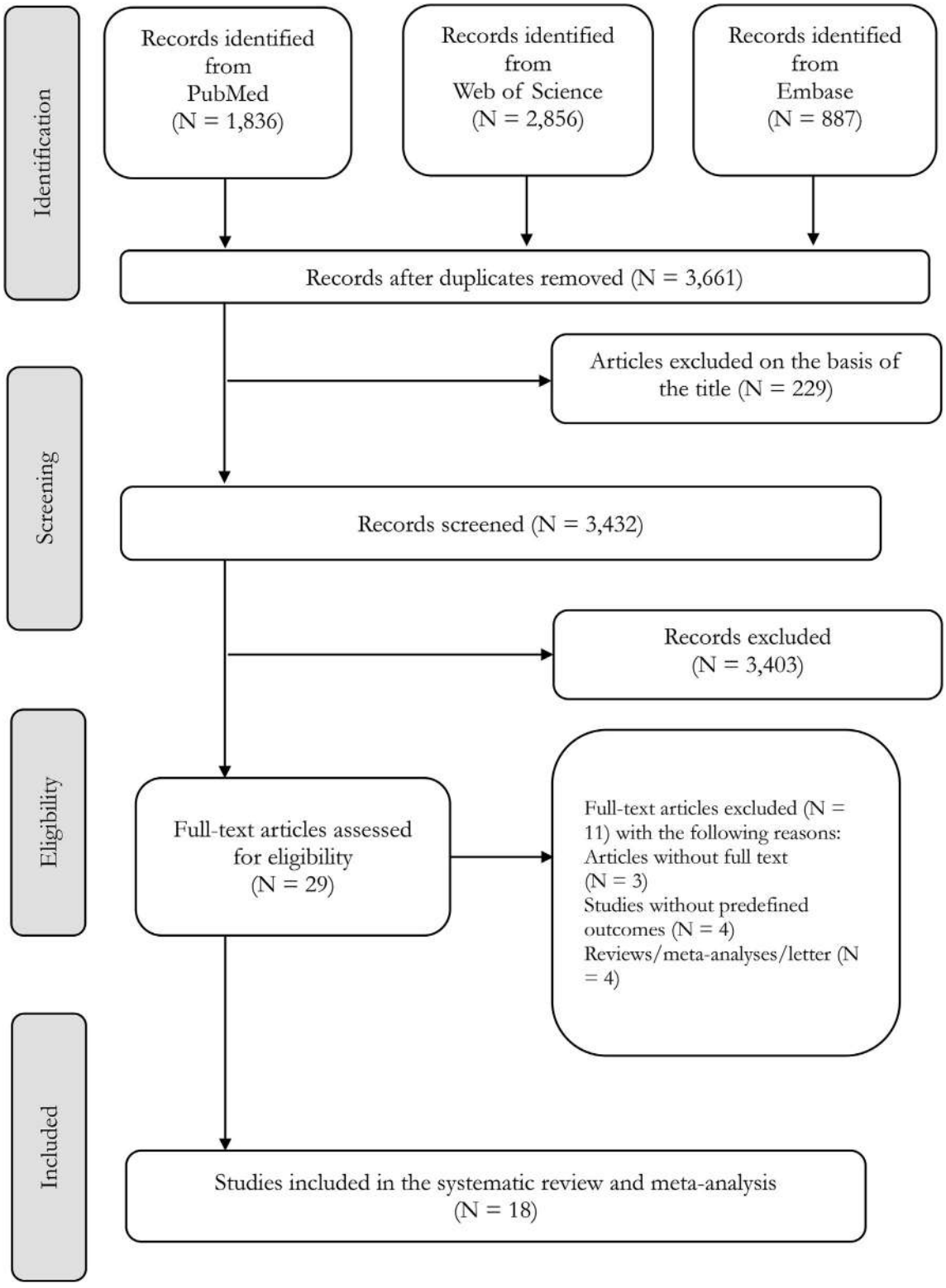


Fig. 1. Flow diagram of study selection. Identification Screening Eligibility Included Records.

Pre- and post-birth air pollution and telomere length in offspring

A summary of the results of our meta-analyses was presented in Table 2, and the majority of the included studies evaluated the relationships for several exposure windows. Prenatal exposure to $PM_{2.5}$ during the entire pregnancy was related to shortening telomere length in offspring ($\beta = -0.043$, 95% CI: −0.067, −0.018). NO_2 exposure was also associated with shortening telomere length in the 1st trimester ($\beta = -0.016$, 95% CI: −0.027, −0.005) and the 3rd trimester ($\beta = -0.036$, 95% CI: −0.068, −0.003). In all exposure

Table 1
Characteristics of epidemiological studies investigated the associations of prenatal and childhood air pollution exposure with telomere length in children.

Authors, year (Country)	Study design, population, and statistical methods	Exposure assessment	Outcome assessment	Adjusted confounders	Main results
Prenatal exposure					
Song et al., 2019 (China)	Birth cohort 743 mother-new born pairs from a birth cohort in Wuhan, China Multiple linear regression models	**Exposure variable:** $PM_{2.5}$, PM_{10}, SO_2, CO, and NO_2 **Exposure assessment method:** Spatial−temporal land use regression models **Windows of exposure:** Entire pregnancy, 1st trimester, 2nd trimester, and 3rd trimester	Telomere length, measured from cord blood DNA using a quantitative real-time polymerase chain reaction (qRT-PCR) method	Maternal age, pre-pregnancy BMI, parity, educational level, passive smoking during pregnancy, gestational diabetes, hypertensive disorders of pregnancy, infant sex, birth weight, gestational age, and season at delivery.	**Entire pregnancy** $PM_{2.5}$: β = −3.54 (−7.34, 0.44) PM_{10}: β = −1.86 (−4.82, 1.20) SO_2: β = −16.82 (−29.86, −1.37) NO_2: β = −1.52 (−6.65, 3.89) CO: β = −1.07 (−4.94, 2.95) **1st trimester** $PM_{2.5}$: β = 0.04 (−1.72, 1.83) PM_{10}: β = 0.34 (−1.33, 2.09) SO_2: β = −0.41 (−6.12, 5.65) NO_2: β = −1.20 (−4.75, 2.48) CO: β = 0.50 (−1.14, 2.17) **2nd trimester** $PM_{2.5}$: β = −0.46 (−2.82, 1.97) PM_{10}: β = −0.16 (−2.27, 2.00) SO_2: β = −2.50 (−10.05, 5.69) NO_2: β = 1.11 (−2.72, 5.09) CO: β = 0.44 (−1.79, 2.74) **3rd trimester** $PM_{2.5}$: β = −3.71 (−6.06, −1.30) PM_{10}: β = −3.24 (−5.29, −1.14) SO_2: β = −11.07 (−18.86, −2.53) NO_2: β = −2.13 (−5.88, 1.78) CO: β = −3.67 (−6.27, −1.00)
Rosa et al., 2019 (Mexico)	Birth cohort 423 women residents in Mexico City and affiliated with the Mexican Social Security System Distributed lag model incorporating weekly averages for $PM_{2.5}$ and Bayesian distributed lag interaction models	**Exposure variable:** $PM_{2.5}$ **Exposure assessment method:** Validated satellite-based spatiotemporally resolved prediction model **Windows of exposure:** Entire pregnancy, 1st trimester, 2nd trimester, and 3rd trimester	Leukocyte telomere length in cord blood was measured using the qPCR method	Sex, maternal age at delivery, prenatal exposure to environmental tobacco smoke, pre-pregnancy BMI, gestational age, birth season, and batch	**1st trimester** $PM_{2.5}$: β = −6.76 (−14.78, 1.01) **2nd trimester** $PM_{2.5}$: β = 6.18 (−5.82, 20.92) **3rd trimester** $PM_{2.5}$: β = 0 (−6.76, 8.33) **Entire pregnancy** $PM_{2.5}$: β = −0.99 (−11.31, 11.63)
Martens et al., 2017 (Belgium)	Cohort 641 mother −newborn pairs from ENVIRONAGE Distributed lag models	**Exposure variable:** $PM_{2.5}$ **Exposure assessment method:** High-resolution spatial −temporal interpolation model (kriging) in combination with a dispersion model **Windows of exposure:** Entire pregnancy, 1st trimester, 2nd trimester, and 3rd trimester	Telomere length from cord blood and placental tissue was measured using the qPCR method	Date of delivery, gestational age, maternal body mass index, maternal age, paternal age, newborn sex, newborn ethnicity, season of delivery, parity, maternal smoking status, maternal educational level, pregnancy complications, and ambient temperature	**Cord blood** Entire pregnancy: β = −8.4 (−13.5, −2.9) 1st trimester: β = −0.8 (−4.7, 3.2) 2nd trimester: β = −9.8 (−13.3, −6.2) 3rd trimester: β = 2.6 (−1.4, 6.8) **Placental tissue** Entire pregnancy: β = −12.5 (−18.4, −6.2) 1st trimester: β = −0.8 (−5.5, 4.1) 2nd trimester: β = −7.4 (−11.7, −2.9) 3rd trimester: β = −4.5 (−9.0, 0.2)

Clemente et al., 2019 (Europe)	Cohort 1396 mother−child pairs recruited from HELIX Generalized additive models and multiple linear mixed models	**Exposure variable:** $PM_{2.5}$ and NO_2 **Exposure assessment method:** LUR and dispersion models **Windows of exposure:** Entire pregnancy, 1st trimester, 2nd trimester, and 3rd trimester	Telomere length, measured from blood (buffy coat) using modified qRT-PCR	Child's age, sex, qRT-PCR batch, maternal age, maternal education, maternal smoking status during pregnancy, child ethnicity, child BMI, and parental smoking at eight y	**Entire pregnancy** NO_2: β = −1.5 (−2.8, −0.2) $PM_{2.5}$: β = −0.7 (−2.0, 0.6) **1st trimester** NO_2: β = −1.6 (−2.8, 0.4) $PM_{2.5}$: β = −0.8 (−2.3, 0.7) **2nd trimester:** NO_2: β = −1.3 (−2.6, −0.04) $PM_{2.5}$: β = −0.1 (−1.3, 1.1) **3rd trimester:** NO_2: β = −1.6 (−2.9, −0.4) $PM_{2.5}$: β = −0.3 (−1.5, 0.8)
Lee et al., 2020 (USA)	Cohort 155 mothers recruited from the Programming of Intergenerational Stress Mechanisms (PEISM) study Bayesian distributed lag interaction models	**Exposure variable:** $PM_{2.5}$ **Exposure assessment method:** Validated spatiotemporally resolved satellite-based model **Windows of exposure:** Entire pregnancy, 1st trimester, 2nd trimester, and 3rd trimester	Telomere length from cord blood, measured by a qRT-PCR	Maternal age, self-reported ethnicity, marital status, education level, maternal lifetime stress, antioxidant intake and infant sex	**Entire pregnancy** β = −0.29 (−0.49, −0.10)
Scholten et al., 2021 (Denmark)	Cohort 296 mothers recruited from the project Maternal Stress and Placental Function The distributed lag models and multivariate linear regression models	**Exposure variable:** $PM_{2.5}$, PM_{10}, SO_2, CO, NO_2, NO_x **Exposure assessment method:** The high-resolution and spatial−temporal air pollution modeling system DEHM-UBM-AirGIS **Windows of exposure:** three trimesters of the pregnancy (weeks 1−12, weeks 13−26, weeks 27−40) and for the overall pregnancy (weeks 1−40) as well as for the period prior to the estimated conception (weeks − eight to 0)	Telomere length from umbilical cord blood, placenta tissue and maternal blood samples, measured by a qRT-PCR	Maternal age, maternal BMI, maternal educational level, maternal smoking habit, ambient temperature, indoor exposure, newborn gender, season of delivery, gestational age in days, pregnancy complications, parity, mode of delivery, newborn weight, length, and circumference of head	**Umbilical cord blood cells:** **Prior to conception** $PM_{2.5}$: β = 4 (−8, 18) PM_{10}: β = 7 (−10, 27) NO_2: β = 11 (2, 22) NO_x: β = 8 (−2, 18) CO: β = 0 (−25, 18) SO_2: β = 18 (4, 33) **Entire pregnancy** $PM_{2.5}$: β = 11 (−9, 36) PM_{10}: β = 1 (−20, 28) NO_2: β = 9 (−6, 27) NO_x: β = 9 (−5, 25) CO: β = 28 (−7, 78) SO_2: β = −20 (−38, 3) **1st trimester** $PM_{2.5}$: β = 7 (−10, 27) PM_{10}: β = −10 (−29, 15) NO_2: β = −5 (−17, 7) NO_x: β = −12 (−22, 0) CO: β = −13 (−33, 13) SO_2: β = 58 (25, 98) **2nd trimester** $PM_{2.5}$: β = 18 (−5, 46) PM_{10}: β = 12 (−16, 49) NO_2: β = 20 (3, 39) NO_x: β = 19 (3, 37) CO: β = 70 (24, 132) SO_2: β = −36 (−52, −25) **3rd trimester** $PM_{2.5}$: β = −23 (−35, −9) PM_{10}: β = 7 (−27, 19) NO_2: β = −20 (−31, −6) NO_x: β = −9 (−22, 6)

(*continued on next page*)

Table 1 (*continued*)

Authors, year (Country)	Study design, population, and statistical methods	Exposure assessment	Outcome assessment	Adjusted confounders	Main results
					CO: β = −29 (−48, −5) SO_2: β = −33 (−47, −16)
Mandakh et al., 2021 (Sweden)	Cross-sectional 42 preeclamptic and 95 arbitrarily selected normotensive pregnant women with gestational ambient NO_x exposure assessment Linear and logistic regression models	**Exposure variable:** NO_x **Exposure assessment method:** Gaussian dispersion model **Windows of exposure:** Entire pregnancy, 1st trimester, 2nd trimester, and 3rd trimester	To quantify telomere length from placental, qRT-PCR were performed	Maternal age, pregestational BMI, parity, gestational age, season of birth and fetal sex	**Entire pregnancy** High vs low: β = 0.08 (−0.06, 0.21) **1st trimester** High vs low: β = −0.03 (−0.11, 0.17) **2nd trimester** High vs low: β = 0.07 (−0.08, 0.21) **3rd trimester** High vs low: β = 0.07 (−0.06, 021)
Isaevska et al., 2022 (Italy)	Cohort PM_{10} daily exposure levels, based on maternal residential address, were estimated for different gestational periods using models based on satellite data Distributed lag models	**Exposure variable:** PM_{10} **Exposure assessment method:** Daily PM_{10} concentrations were estimated at 1-km^2 grid using the Random Forest (RF) method **Windows of exposure:** Each week, 1st trimester, 2nd trimester, and 3rd trimester, entire pregnancy	The average relative TL was measured using the monochrome multiplex quantitative PCR (MMq-PCR)	Study center, maternal education, maternal pre-pregnancy BMI, parity, smoking in pregnancy, child's sex, PM_{10} concentrations, gestational age, season of birth, pregnancy complications	**1st trimester** PM_{10}: β = −0.0157 (−0.0358, 0.0044) **2nd trimester** PM_{10}: β = 0.0019 (−0.0131, 0.0169) **3rd trimester** PM_{10}: β = 0.0143 (−0.0035, 0.0320) **Entire pregnancy:** PM_{10}: β = 0.0024 (−0.0296, 0.0344)
Durham et al., 2022 (USA)	Cohort 197 pairs of Dominican and African American mother-child Multivariable linear regression models	**Exposure variable:** $PM_{2.5}$ **Exposure assessment method:** To assign estimates of fine particulate matter ($PM_{2.5}$), the validated spatio-temporal air pollution exposure models was utilized **Windows of exposure:** 1st trimester, 2nd trimester, and 3rd trimester	Telomere length from umbilical cord blood was measured using the qPCR method	Gestational age at birth, child sex, ethnicity, maternal age, and maternal education	**1st trimester** $PM_{2.5}$: β = 0.039 (−0.039, 0.117) **2nd trimester** $PM_{2.5}$: β = −0.037 (−0.114, 0.039) **3rd trimester** $PM_{2.5}$: β = 0.042 (−0.036, 0.120) **Entire pregnancy** $PM_{2.5}$: β = 0.063 (−0.021, 0.147)
Childhood exposure					
Walton et al., 2016 (UK)	Cross-sectional 333 children aged 8 −9 years in 23 schools in east London Linear mixed-effects models	**Exposure variable:** NO_x, NO_2, PM_{10}, and $PM_{2.5}$ **Exposure assessment method:** Kings College London, UK urban models **Windows of exposure:** Annual air pollution exposure, exposure over previous week,	Telomere length, measured from oral DNA using the MMq-PCR	Age, sex, ethnicity, study year, IgA, cortisol and included a random intercept for school	**Annual air pollution exposure** NO_x: β = 1.004 (1.002, 1.006) NO_2: β = 1.012 (1.005, 1.016) $PM_{2.5}$: β = 1.116 (1.056, 1.179) PM_{10}: β = 1.047 (1.024, 1.071) **Exposure over previous week** NO_x: β = 1.003 (1.000, 1.006) NO_2: β = 1.007 (1.000, 1.014) $PM_{2.5}$: β = 1.013 (1.000, 1.025) PM_{10}: β = 1.010 (1.002, 1.018) **Exposure over previous day**

		and exposure over previous day			NO_x: β = 1.001 (0.999, 1.002) NO_2: β = 1.001 (0.997, 1.006) $PM_{2.5}$: β = 1.004 (0.995, 1.014) PM_{10}: β = 1.003 (0.997, 1.009)
Moslem et al., 2020 (Iran)	Cross-sectional 200 preschool children (5–7 years old) recruited from 27 kindergartens in Sabzevar, Iran Mixed linear regression models	**Exposure variable:** $PM_{2.5}$, PM_{10} **Exposure assessment method:** LUR models **Windows of exposure:** Annual air pollution exposure of 2017	Telomere length, measured from blood using qRT-PCR	Age, sex, BMI, parental education, income, tobacco exposure at home, illiterate percent per census tract, and unemployed percent per census tract	**Annual air pollution exposure** **Home** $PM_{2.5}$: β = −0.15 (−0.23, −0.07) PM_{10}: β = −0.13 (−0.20, −0.06) **Kindergarten-outdoor** $PM_{2.5}$: β = −0.23 (−0.39, −0.08) PM_{10}: β = −0.13 (−0.20, −0.06) **Kindergarten-indoor** $PM_{2.5}$: β = −0.18 (−0.36, −0.01) PM_{10}: β = −0.11 (−0.24, 0.01)
Clemente et al., 2019 (Europe)	Cohort 1396 mother–child pairs recruited from HELIX Generalized additive models and multiple linear mixed models	**Exposure variable:** $PM_{2.5}$, NO_2 **Exposure assessment method:** LUR and dispersion models **Windows of exposure:** One-year childhood air pollution exposure before the telomere length measurements	Telomere length, measured from blood (buffy coat) using modified qRT-PCR	Child's age, sex, qRT-PCR batch, maternal age, maternal education, maternal smoking status during pregnancy, child ethnicity, child BMI, and parental smoking at eight y	**Annual air pollution exposure** NO_2: β = −1.6 (−2.9, −0.4) $PM_{2.5}$: β = −1.4 (−2.9, 0.1)
Hautekiet et al., 2021 (Belgium)	Cohort 197 primary school children Mixed-effects models	**Exposure variable:** $PM_{2.5}$, NO_2 **Exposure assessment method:** Validated spatiotemporally resolved satellite-based model **Windows of exposure:** Recent (week) exposure, sub-chronic exposure (month) and chronic (year) exposure	Telomere length, measured from oral cells using qRT-PCR	Sex, age, BMI, socioeconomic status, passive smoking, season of examination, examination, apparent temperature and the random effect of school and subject	**Recent (day)** $PM_{2.5}$: β = −4.7 (−9.5, 0.4) NO_2: β = −8.9 (−15.4, −1.9) **Sub-chronic (month)** $PM_{2.5}$: β = −6.2 (−10.6, −1.6) NO_2: β = −5.0 (−10.7, 1.0) **Chronic (year)** $PM_{2.5}$: β = −6.4 (−11.8, −0.7) NO_2: β = −0.9 (−5.0, 3.4)
Adulthood exposure					
Wong et al., 2014 (USA)	Panel 48 male boilermaker workers Linear mixed-effects regression models with random intercept	**Exposure variable:** $PM_{2.5}$ **Exposure assessment method:** 38 independent area $PM_{2.5}$ measures were obtained using KTL Cyclones from an oil-burning power plant located in eastern Massachusetts **Windows of exposure:** Month, year, and career prior to each blood draw	Telomere length, measured from leukocyte using qRT-PCR	Neutrophil, lymphocyte, monocyte, eosinophil, current smoking intensity, age at baseline blood draw, body mass index, and years as a boilermaker	**Career prior to blood draws** $PM_{2.5}$: β = −0.021 (−0.048 0.006) **Year prior to blood draws** $PM_{2.5}$: β = −0.002 (−0.053, 0.009) **Month prior to blood draws** $PM_{2.5}$: β = −0.040 (−0.080, 0.001)
Xia et al., 2015 (China)	Panel 35 patients with type two diabetes Linear mixed-effect models	**Exposure variable:** $PM_{2.5}$, SO_2, NO_2, CO, O_3 **Exposure assessment method** PM were measured on	Telomere length, measured from blood using qRT-PCR	Age, sex, body mass index, education status, annual income per capita, medication use, history of diabetes	24-h mean air pollution $PM_{2.5}$: 0.11 (−0.97, 1.19) SO_2: 0.48 (−1.01, 1.97) NO_2: 0.25 (−0.76, 1.26)

(*continued on next page*)

Table 1 (*continued*)

Authors, year (Country)	Study design, population, and statistical methods	Exposure assessment	Outcome assessment	Adjusted confounders	Main results
		the rooftop of the building (about 18 m high) of the Tianping Community Health Center (TCHC) using the Environmental Dust Monitor 365. Hourly concentrations of gaseous pollutants, including SO_2, NO_2 and CO, were derived from a fixed-site nation-controlled station that was about 2.5 km away from the TCHC. **Windows of exposure:** Lag o day, lag 1 day, lag 2 day, lag 3 day, lag 4–7 day			CO: 0.02 (−1.16, 1.20) O_3: 0.28 (−0.76, 1.32)
Pieters et al., 2016 (Belgium)	Cross-sectional 166 non-smokers elderly participants Multi-variable linear regression models	**Exposure variable:** $PM_{2.5}$ **Exposure assessment method** The annual exposure levels of $PM_{2.5}$ were estimated for each participant's home address using a high-resolution spatial interpolation method (kriging method) **Windows of exposure:** Last year, last month, and last week	Telomere length was measured by the qRT-PCR method developed by Cawthon with minor adaptations	Sex, age, BMI, socioeconomic status, statin use, past smoking status, white blood cell count, and percentage of neutrophils	**Last year** $PM_{2.5}$: β = −0.040 (−0.065,3–0.017) **Last month** $PM_{2.5}$: β = 0.029 (0.017, 0.041) **Last week** $PM_{2.5}$: β = 0.003 (−0.007, 0.014)
Carugno et al., 2021 (Italy)	Cross-sectional 1792 participants with overweight/obesity Multi-variable linear regression models	**Exposure variable:** PM_{10} **Exposure assessment method:** Daily PM_{10} concentration series from air quality monitors and daily PM_{10} concentrations estimated with the Flexible Air quality Regional Model **Windows of exposure:** Daily lags, lag 0–1 to lag 0–30	Telomere length, measured from blood sample using qRT-PCR	Sex, age, education, BMI, alcohol consumption, pack-years of smoking, place of living, type two diabetes, metabolic syndrome, cancer	**Lag 0** PM_{10}: β = −0.0051 (−0.0098, −0.0005) **Annual mean** PM_{10}: β = −0.0257 (−0.0506 −0.0008)
Dioni et al., 2011 (Italy)	Panel 63 male workers free of cardiopulmonary disease or cancer	**Exposure variable:** PM_{10} **Exposure assessment method:** PM_{10} was measured	Leukocyte telomere length was measured using the qRT-PCR method developed by Cawthon with minor adaptations	Age, BMI, pack-years, and percent lymphocytes	**Baseline** PM_{10}: β = 0.08 (−0.03, 0.19) **Post exposure** PM_{10}: β = 0.30 (0.11, 0.49)

	Linear regression models	during the 3 days between the baseline and post exposure blood drawing using a GRIMM 1100 light-scattering dust analyzer **Windows of exposure:** Baseline, postexposure, and difference			**Difference** PM_{10}: $\beta = 0.23$ (0.08, 0.38)
Hou et al., 2012 (China)	Panel 120 truck drivers and 120 office workers Mixed-effect regression models	**Exposure variable:** $PM_{2.5}$, PM_{10} **Exposure assessment method:** Personal $PM_{2.5}$ and Elemental Carbon were measured using light-weight monitors. Ambient PM_{10} was obtained from local monitoring stations **Windows of exposure:** Examination day (24-h average), 1-day mean (24- hour average of the day before the examinations), as well as averages of the 24 h means of 1−2 days, 1−5 days, 1−7 days, 1−10 days and 1−14 days before the examinations	Blood telomere length was measured using the qRT-PCR method developed by Cawthon with minor adaptations	Age, sex, BMI, number of cigarettes smoked during examination time, day of the week, usage of central heating, time used for commuting to work, temperature, and dew point	**During the eight work hours** $PM_{2.5}$: $\beta = 0.052$ (0.015, 0.091) **Examination days:** PM_{10}: $\beta = 0.077$ (0.037, 0.119)

windows, there were no significant associations between prenatal SO_2 and CO exposures and telomere length. In terms of childhood air pollution exposure, $PM_{2.5}$, PM_{10}, and NO_2 exposures were not associated with telomere length in children.

The forest plots of the associations between prenatal and childhood air pollution exposure and telomere length in children were presented in Supplementary Material, Figs. S1–S8. In this study, subgroup analysis was not performed due to the small number of studies included in each air pollutant and telomere length group stratified by exposure window.

Air pollution and telomere length in adults

A total of six studies was included that regarding the association of air pollution exposure and telomere length. There were three studies that examined the association between PM_{10} and telomere length, including one based on long-term[44] and three based on short-term exposure.[44–46] Four studies examined the relationship between $PM_{2.5}$ and telomere length, of which two evaluated long-term relationships[23,24] three examined short-term relationships.[24,25,46] According to the meta-analyses, long-term exposure to $PM_{2.5}$ is associated with shortening telomere length, while short-term exposure to $PM_{2.5}$ and PM_{10} does not have any impact on telomere length. The forest plots of adulthood air pollution exposure with telomere length were presented in Supplementary Material, Figs. S9–S10.

Publication bias

All funnel plots of prenatal $PM_{2.5}$ exposure in each exposure window were nearly symmetrical (Supplementary Material, Figs. S11–S14), which indicated the absence of publication bias. We also performed the Egger's test to quantitively assess publication bias when three or more studies were available. The *P*-values of the Egger's tests for most air pollutants exposure in different windows were more than 0.05 (Table 2), which suggested no publication bias was detected among the meta-analytic exposure-outcome combinations. However, it should be noted that the *P*-value of Egger's test for prenatal NO_2 exposure during the 2nd trimester was 0.025, which indicated that publication bias was present. We thus used the trim-and-fill method to adjust for potential publication bias and the result was noted in Supplementary Material, Fig. S15.

Sensitivity analysis

For the sensitivity analysis of prenatal $PM_{2.5}$ exposure, we used telomere length data from placental blood cells instead of cord blood cells. After replaced the estimates, the summary estimates were −0.036 (−0.053, −0.019), −0.008 (−0.023, 0.007), −0.025 (−0.066, 0.015), and −0.031 (−0.072, 0.010) for the entire pregnancy, 1st, 2nd, and 3rd trimester, respectively. The results were consistent with each other, which indicate the robustness of meta-analytic results.

For childhood air pollution exposure, we replaced the effect estimates using data from kindergarten-outdoor instead of home, the summary estimates were −0.043 (−0.161, 0.074) for $PM_{2.5}$ and 0.186 (−0.345, 0.717) for PM_{10}, respectively. In addition, we also replaced the estimates by using data from kindergarten-indoor instead of the data from home, the summary estimates were −0.048 (−0.174, 0.079) for $PM_{2.5}$ and 0.193 (−0.324, 0.710). The pooled estimates of $PM_{2.5}$ and PM_{10} exposures in childhood were robust.

To evaluate the stability of results, leave-one-out method sensitivity analysis was used for adulthood air pollution exposure. There was no significant change in the pooled estimates of $PM_{2.5}$ exposure when each study was excluded except for PM_{10} exposure, indicated that the result of $PM_{2.5}$ exposure was generally robust (Supplementary Material, Figs. S16–18).

Confidence in the body of evidence

This meta-analysis included observational studies, which resulted in an initial rating of 'moderate confidence' according to the NTP/OHAT framework. Five factors could downgrade the rating. Inconsistency was not the main downgrading factor since there was not substantial heterogeneity across the significant meta-analytic exposure-outcome combinations. As most studies were classified as Tier 1, there was no decrease in confidence regarding the risk of bias. Other factors seem not to upgrade or downgrade the confidence due to the absence of a clear criteria to judge. Overall, the confidence in the body evidence for was moderate for the significant meta-analytic exposure-outcome combinations. Thus, the level of evidence for a significant association between air pollution exposure and telomere length (entire pregnancy $PM_{2.5}$ exposure, the 1st trimester NO_2 exposure, and adulthood $PM_{2.5}$ exposure) was translated to 'moderate'.

Discussion

It is well known that air pollution is a serious concern for public health, which is related to a wide range of health risks for both children and adults. Miri et al. have previously conducted a meta-analysis on the relationship between air pollution exposure and telomere length in adults; however, they included a small number of relevant studies and did not focus on the effect of air pollution exposure on telomere length in children.[30] Isaekska et al. reviewed the relevant studies and concluded that ambient air pollution during the early life was associated with global and locus-specific DNA methylation changes and telomere length shortening, but they did not perform further quantitative analyses.[48] Currently, there is no relevant meta-analysis that addresses the association between early life air pollution and telomere length in offspring. It has been shown that the number of studies investigating the relationships between pre- and post-birth air pollution exposure and telomere length in offspring has increased in recent years, of which some have reported inconsistent results. We conducted a meta-analysis to assess the association between pre- and post-birth air pollution exposure and telomere length in offspring and updated the evidence on air pollution and telomere length in adults in order to comprehensively assess these associations in a systematic and transparent manner.

Generally, this study suggested that $PM_{2.5}$ and NO_2 exposure in prenatal period were related to shortening telomere length in offspring. At the same time, long-term $PM_{2.5}$ exposure during adulthood was associated with telomere length in adults. Nevertheless, these results should be interpreted with caution due to the relatively small number of included studies. Of the positive associations between prenatal air pollutants exposure and telomere length, it can be explained by the truth that fetus telomeres may be variable in response to air pollution exposure during fetal period, while for adult exposure, these short-term effects were not obvious.

Despite the fact that the relevant mechanisms are not fully understood, oxidative stress and inflammation are two of the main mechanisms through which air pollution might eventually damage telomeres.[49] It is known that reactive oxygen species are capable of accumulating single-strand nicks in telomeres, which are less readily repaired than other genomic regions.[50] Additionally, air pollution exposure may lead to increased replication speed of leukocytes, accelerating the shortening of telomere length.[51]

Table 2
Summary estimates and 95% confidence intervals (CIs) in random-effects meta-analysis for the associations of air pollution exposure with telomere length.

Air pollutant exposure and exposure window	Studies included[a]	Sample size	Summary estimate (95% CI)[b]	Heterogeneity				*P*-value of Egger's test
				Q	*P*-value	I^2 (%)	τ^2	
$PM_{2.5}$ (prenatal)								
Entire pregnancy	1, 2, 3, 4, 5, 6, 9	4381	−0.043 (−0.067, −0.018)	9.55	0.145	37.1	0.0003	0.581
Entire pregnancy[c]	1, 2, 3, 4, 5, 6, 9	4381	−0.036 (−0.053, −0.019)	6.46	0.373	7.1	<0.0001	0.606
1st trimester	1, 2, 3, 4, 6, 9	4226	−0.008 (−0.023, 0.006)	4.87	0.432	0.0	<0.0001	0.375
1st trimester[c]	1, 2, 3, 4, 6, 9	4226	−0.008 (−0.023, 0.007)	4.84	0.436	0.0	<0.0001	0.402
2nd trimester	1, 2, 3, 4, 6, 9	4226	−0.037 (−0.093, 0.019)	30.20	<0.001	83.4	0.0032	0.875
2nd trimester[c]	1, 2, 3, 4, 6, 9	4226	−0.025 (−0.066, 0.015)	14.74	0.012	66.1	0.0013	0.973
3rd trimester	1, 2, 3, 4, 6, 9	4226	−0.013 (−0.055, 0.028)	18.42	0.004	72.9	0.0016	0.763
3rd trimester[c]	1, 2, 3, 4, 6, 9	4226	−0.031 (−0.072, 0.010)	17.36	0.004	71.2	0.0015	0.306
PM_{10} (prenatal)								
Entire pregnancy	1, 5, 8	1625	−0.034 (−0.074, 0.005)	8.68	0.013	77.0	0.0009	0.757
1st trimester	1, 5, 8	1625	−0.001 (−0.003, 0.005)	1.75	0.416	0.0	<0.0001	0.545
2nd trimester	1, 5, 8	1625	−0.001 (−0.021, 0.018)	0.53	0.766	0.0	<0.0001	0.326
3rd trimester	1, 5, 8	1625	−0.015 (−0.050, 0.021)	10.02	0.007	80.0	0.0006	0.658
NO_2 (prenatal, including NO_x scaled)								
Entire pregnancy	1, 4, 5, 5 (NO_x)	2294	−0.009 (−0.023, 0.005)	3.11	0.375	3.6	<0.0001	0.206
1st trimester	1, 4, 5, 5 (NO_x)	2294	−0.016 (−0.027, −0.005)	1.79	0.616	0.0	<0.0001	0.249
2nd trimester	1, 4, 5, 5 (NO_x)	2294	−0.021 (−0.025, 0.066)	8.96	0.030	66.5	0.0011	0.025
3rd trimester	1, 4, 5, 5 (NO_x)	2294	−0.036 (−0.068, −0.003)	17.42	0.001	82.8	0.0007	0.180
SO_2 (prenatal)								
Entire pregnancy	1, 5	1039	−0.110 (−0.321, 0.101)	2.14	0.143	53.3	0.0156	Not available
1st trimester	1, 5	1039	−0.371 (−0.438, 1.180)	9.72	0.002	89.7	0.3088	Not available
2nd trimester	1, 5	1039	−0.245 (−0.265, 0.755)	27.70	<0.001	96.4	0.1308	Not available
3rd trimester	1, 5	1039	−0.236 (−0.669, 0.197)	15.20	<0.001	93.4	0.0915	Not available
CO (prenatal)								
Entire pregnancy	1, 5	898	−0.076 (−0.245, 0.396)	1.75	0.185	43.0	0.0338	Not available
1st trimester	1, 5	898	−0.087 (−0.267, 0.094)	102.04	<0.001	99.0	0.0168	Not available
2nd trimester	1, 5	898	−0.410 (−0.520, 1.340)	6.39	0.011	84.4	0.3896	Not available
3rd trimester	1, 5	898	−0.187 (−0.538, 0.164)	5.72	0.017	82.5	0.0545	Not available
$PM_{2.5}$ (childhood)								
One year exposure	4, 10, 11, 12	2699	−0.051 (−0.172, 0.071)	34.60	<0.001	91.3	0.0118	0.709
One year exposure[d]	4, 10, 11, 12	2699	−0.043 (−0.161, 0.074)	30.90	<0.001	90.3	0.0109	0.820
One year exposure[e]	4, 10, 11, 12	2699	−0.048 (−0.174, 0.079)	27.65	<0.001	89.2	0.0126	0.567
Previous week exposure	10, 12	1103	−0.041 (−0.409, 0.327)	6.55	0.010	84.7	0.0603	Not available
PM_{10} (childhood)								
One year exposure	10, 11	533	−0.188 (−0.340, 0.716)	19.24	<0.001	94.8	0.1378	Not available
One year exposure[d]	10, 11	533	−0.186 (−0.345, 0.717)	19.90	<0.001	95.0	0.1396	Not available
One year exposure[e]	10, 11	533	−0.193 (−0.324, 0.710)	17.98	<0.001	94.4	0.1317	Not available
NO_2 (childhood, including NO_x scaled)								
One year exposure	4, 10, 10 (NO_x), 12	2499	−0.030 (−0.015, 0.074)	40.08	<0.001	92.5	0.0018	0.283
Previous week exposure	10, 10 (NO_x), 12	1103	−0.011 (−0.057, 0.079)	11.16	0.004	82.1	0.0029	0.870
$PM_{2.5}$ (adulthood)								
Long-term exposure	13, 15	214	−0.045 (−0.084, −0.005)	0.73	0.392	0.0	<0.0001	Not available
Short-term exposure	14, 15, 18	312	0.007 (−0.004, 0.018)	6.29	0.043	68.2	0.0001	0.449
PM_{10} (adulthood)								
Short-term exposure	16, 17, 18	1975	−0.004 (−0.011, 0.003)	22.43	<0.001	91.1	<0.0001	Not available

[a] Studies included are different in each air pollutant exposure window meta-analysis depending on the data they published. References are as follows: 1. Song et al. (2019); 2. Rosa et al. (2019); 3. Maretens et al. (2017); 4. Clemente et al. (2019); 5. Lee et al. (2020); 6. Scholten et al. (2021); 7. Mandakh et al. (2021); 8. Isaevska et al. (2022); 9. Durham et al. (2022); 10. Walton et al. (2016); 11. Moslem et al. (2020); 12. Hautekiet et al. (2021); 13. Wong et al. (2014); 14. Xia et al. (2015); 15. Pieters et al. (2016); 16. Carugno et al. (2021); 17. Dioni et al. (2011); 18. Hou et al. (2012).
[b] In prenatal and childhood exposure, units are 10 μg/m^{3,} increase in $PM_{2.5}$, PM_{10} and NO_2, 1 ppb increase in SO_2, and 100 μg/m^{3,} increase in CO; in adulthood exposure, unit is 5 μg/m^{3,} increase in $PM_{2.5}$ and PM_{10}.
[c] Sensitivity analysis using telomere length data from placental blood cells instead of cord blood cells.
[d] Sensitivity analysis using air pollution data from kindergarden-outdoor instead of home.
[e] Sensitivity analysis using air pollution data from kindergarden-indoor instead of home.

The observed heterogeneity between studies may be attributed to a wide range of aspects. Here, we discussed the role of exposure assessment among the included studies. Among the included studies on pre- and post-birth air pollution exposure, varied estimate models were used by each individual studies, such as land use regression model,[27,28,43] spatial–temporal model,[29] spatiotemporally resolved satellite-based model,[40] spatiotemporal interpolation model,[18,38] and the multiscale integrated air pollution model.[42] On the other hand, regarding studies on short-term air pollution exposure and telomere length in adults, the included studies used work history questionnaires and area air measures,[23] personal real-time air pollution measurement,[25,45,46] high-resolution spatial–temporal interpolation model,[24] and fixed monitoring stations.[44] All methods have advantages and disadvantages, and individual studies should choose the most suitable model based on the characteristics of the research, including spatial variations of air pollutants, and quality and availability of data. In the case of a panel study, for example, the measurement of personal air pollution in real-time may be an appropriate method of assessing exposure.

We also found that the evidence on the associations of pre- and post-birth air pollution exposure with telomere length in offspring as well as the relationship between adulthood air pollution exposure and telomere length was inconsistent. The meta-analyses suggested that there were several significant associations among children,

while only significant association was found for long-term $PM_{2.5}$ exposure among adults. The possible reasons may include: (1) The studies on adults always investigated occupationally exposed individuals, such as truck drivers[46] steel workers,[45] police traffic officers,[13] and boilermakers[23] with a short-term air pollution exposure. Thus, the possibility of the bias of healthy worker effect cannot be ruled out. (2) Telomere length in newborns varies greatly, and telomere attrition rates are higher during the first four years of life than later in life.[52–54] In addition, fetal telomeres may vary depending on the period during which they are exposed to air pollution.[55,56] Thus, the effects were more pronounced for early life air pollution exposure. Further studies are needed to elucidate the potential effect of air pollution on telomere length in different populations.

In this study, several strengths should be acknowledged. First, this study evaluates the relationships between pre- and post-birth air pollution exposure and telomere length in offspring as well as air pollution and telomere length in adults. Second, this study evaluated the risk of bias in individual studies and the level of confidence in the body of evidence, as opposed to previous similar systematic reviews, which failed to address this issue adequately. Finally, the findings of this study have provided insight into the vulnerability to telomere length in offspring exposed to air pollution.

The study has some limitations, regardless of its strengths. First, it is notable that certain air pollutant and telomere length combinations exhibit high levels of heterogeneity between studies. Differences in study design, the population, study location, the measurement and standardization of telomere length, and the level of confounder adjustments may explain the observed heterogeneity; however, the small number of studies included prohibits any verification of this explanation. Second, each measured air pollutant-telomere length combination contains relatively few studies since this research topic has only recently attracted attention. In this way, we are unable to perform some of the regular analyses of meta-analysis, such as subgroup analyses and meta-regression analyses. Lastly, most of the included studies were conducted in developed countries in Europe and North America, and therefore, we may not be able to generalize our findings to other countries.

Conclusions

This study indicated that the entire pregnancy $PM_{2.5}$ exposure and the 1st trimester NO_2 exposure were related to shortening telomere length in offspring; however, air pollution exposure in childhood was not statistically associated with telomere length in children and short-term air pollution exposure was also not related to telomere length in adults. The results of these studies should be interpreted with caution due to the moderate to high degree of heterogeneity within studies and the relatively small number of studies involving some combinations of air pollutants and telomere length. In order to overcome the limitations of previous studies and elaborate on causal relationships, further studies involving longitudinal data with a standardized telomere measurement method are necessary.

Author statements

Ethical approval

Approval was not required as this analysis was based on published data.

Data availability statement

The raw data supporting the conclusions of this article was shown in the forest plots.

Authorship contributions

Z.Q.Z and S.W.C conceived the idea, performed the statistical analysis, and drafted this meta-analysis. S.Y.G and Y.W selected and searched the relevant papers. Z.Q.Z and C.Y.H assessed each study. Z.Q.Z is the guarantor of the overall content. X.J.Z and C.Y.H supervised the entire study process and contributed to the critical revision of the manuscript. All authors revised and approved the final manuscript.

Funding

This study was supported by the Key Projects of Natural Science Research of Anhui Provincial Department of Education (KJ2020A0163) and National Natural Science Foundation of China (82073565).

Competing interests

None declared.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2022.11.022.

References

1. Juginovic A, Vukovic M, Aranza I, Bilos V. Health impacts of air pollution exposure from 1990 to 2019 in 43 European countries. *Sci Rep* 2021;**11**:22516.
2. Doiron D, de Hoogh K, Probst-Hensch N, Fortier I, Cai Y, De Matteis S, et al. Air pollution, lung function and COPD: results from the population-based UK Biobank study. *Eur Respir J* 2019;**54**.
3. Nhung NTT, Amini H, Schindler C, Kutlar Joss M, Dien TM, Probst-Hensch N, et al. Short-term association between ambient air pollution and pneumonia in children: a systematic review and meta-analysis of time-series and case-crossover studies. *Environ Pollut* 2017;**230**:1000–8.
4. Kim SY, Kim SH, Wee JH, Min C, Han SM, Kim S, et al. Short and long term exposure to air pollution increases the risk of ischemic heart disease. *Sci Rep* 2021;**11**:5108.
5. Wang N, Mengersen K, Tong S, Kimlin M, Zhou M, Wang L, et al. Short-term association between ambient air pollution and lung cancer mortality. *Environ Res* 2019;**179**:108748.
6. Eze IC, Hemkens LG, Bucher HC, Hoffmann B, Schindler C, Kunzli N, et al. Association between ambient air pollution and diabetes mellitus in Europe and North America: systematic review and meta-analysis. *Environ Health Perspect* 2015;**123**:381–9.
7. Huang Y, Zhu M, Ji M, Fan J, Xie J, Wei X, et al. Air pollution, genetic factors, and the risk of lung cancer: a prospective study in the UK biobank. *Am J Respir Crit Care Med* 2021;**204**:817–25.
8. Pedersen M, Giorgis-Allemand L, Bernard C, Aguilera I, Andersen AM, Ballester F, et al. Ambient air pollution and low birthweight: a European cohort study (ESCAPE). *Lancet Respir Med* 2013;**1**:695–704.
9. Qian Z, Liang S, Yang S, Trevathan E, Huang Z, Yang R, et al. Ambient air pollution and preterm birth: a prospective birth cohort study in Wuhan, China. *Int J Hyg Environ Health* 2016;**219**:195–203.
10. Shah AS, Lee KK, McAllister DA, Hunter A, Nair H, Whiteley W, et al. Short term exposure to air pollution and stroke: systematic review and meta-analysis. *BMJ* 2015;**350**:h1295.
11. von Zglinicki T. Oxidative stress shortens telomeres. *Trends Biochem Sci* 2002;**27**:339–44.
12. De Felice B, Nappi C, Zizolfi B, Guida M, Di Spiezio Sardo A, Bifulco G, et al. Telomere shortening in women resident close to waste landfill sites. *Gene* 2012;**500**:101–6.
13. Hoxha M, Dioni L, Bonzini M, Pesatori AC, Fustinoni S, Cavallo D, et al. Association between leukocyte telomere shortening and exposure to traffic pollution: a cross-sectional study on traffic officers and indoor office workers. *Environ Health : a global access science source* 2009;**8**:41.
14. Perera F, Lin CJ, Qu L, Tang D. Shorter telomere length in cord blood associated with prenatal air pollution exposure: benefits of intervention. *Environ Int* 2018;**113**:335–40.
15. Wei H, Zhang C, Silveyra P. The relationships between prenatal smoking exposure and telomere lengths in fetuses, infants, and children: a systematic literature review. *J Addict Nurs* 2020;**31**:243–52.

16. Wu M, Wang L, Song L, Liu B, Liu Y, Bi J, et al. The association between prenatal exposure to thallium and shortened telomere length of newborns. *Chemosphere* 2021;**265**:129025.
17. Coluzzi E, Colamartino M, Cozzi R, Leone S, Meneghini C, O'Callaghan N, et al. Oxidative stress induces persistent telomeric DNA damage responsible for nuclear morphology change in mammalian cells. *PLoS One* 2014;**9**:e110963.
18. Martens DS, Cox B, Janssen BG, Clemente DBP, Gasparrini A, Vanpoucke C, et al. Prenatal air pollution and newborns' predisposition to accelerated biological aging. *JAMA Pediatr* 2017;**171**:1160–7.
19. McCracken J, Baccarelli A, Hoxha M, Dioni L, Melly S, Coull B, et al. Annual ambient black carbon associated with shorter telomeres in elderly men: veterans Affairs Normative Aging Study. *Environ Health Perspect* 2010;**118**:1564–70.
20. Heidinger BJ, Blount JD, Boner W, Griffiths K, Metcalfe NB, Monaghan P. Telomere length in early life predicts lifespan. *Proc Natl Acad Sci U S A* 2012;**109**:1743–8.
21. Bijnens EM, Zeegers MP, Derom C, Martens DS, Gielen M, Hageman GJ, et al. Telomere tracking from birth to adulthood and residential traffic exposure. *BMC Med* 2017;**15**:205.
22. Benetos A, Verhulst S, Labat C, Lai TP, Girerd N, Toupance S, et al. Telomere length tracking in children and their parents: implications for adult onset diseases. *Faseb J* 2019;**33**:14248–53.
23. Wong JY, De Vivo I, Lin X, Christiani DC. Cumulative PM(2.5) exposure and telomere length in workers exposed to welding fumes. *J Toxicol Environ Health, Part A* 2014;**77**:441–55.
24. Pieters N, Janssen BG, Dewitte H, Cox B, Cuypers A, Lefebvre W, et al. Biomolecular markers within the core Axis of aging and particulate air pollution exposure in the elderly: a cross-sectional study. *Environ Health Perspect* 2016;**124**:943–50.
25. Xia Y, Chen R, Wang C, Cai J, Wang L, Zhao Z, et al. Ambient air pollution, blood mitochondrial DNA copy number and telomere length in a panel of diabetes patients. *Inhal Toxicol* 2015;**27**:481–7.
26. Walton RT, Mudway IS, Dundas I, Marlin N, Koh LC, Aitlhadj L, et al. Air pollution, ethnicity and telomere length in east London schoolchildren: an observational study. *Environ Int* 2016;**96**:41–7.
27. Moslem A, Rad A, de Prado Bert P, Alahabadi A, Ebrahimi Aval H, Miri M, et al. Association of exposure to air pollution and telomere length in preschool children. *Sci Total Environ* 2020;**722**:137933.
28. Clemente DBP, Vrijheid M, Martens DS, Bustamante M, Chatzi L, Danileviciute A, et al. Prenatal and childhood traffic-related air pollution exposure and telomere length in European children: the HELIX project. *Environ Health Perspect* 2019;**127**:87001.
29. Rosa MJ, Hsu HL, Just AC, Brennan KJ, Bloomquist T, Kloog I, et al. Association between prenatal particulate air pollution exposure and telomere length in cord blood: effect modification by fetal sex. *Environ Res* 2019;**172**:495–501.
30. Miri M, Nazarzadeh M, Alahabadi A, Ehrampoush MH, Rad A, Lotfi MH, et al. Air pollution and telomere length in adults: a systematic review and meta-analysis of observational studies. *Environ Pollut* 2019;**244**:636–47.
31. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Br Med J* 2009;**339**:b2535.
32. Epa U. *Vocabulary catalog list detail—integrated risk information system (IRIS) 391 glossary*. 2011.
33. Guyatt GH, Oxman AD, Schunemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *J Clin Epidemiol* 2011;**64**:380–2.
34. Schunemann H, Hill S, Guyatt G, Akl EA, Ahmed F. The GRADE approach and Bradford Hill's criteria for causation. *J Epidemiol Community Health* 2011;**65**:392–5.
35. NTP/OHAT (National Toxicology Program/Office of Health Assessment Translation). *Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration - march 4. 2019. Office of health assessment and translation*. U.S. Dept. of Health and Human Services, National Toxicology Program; 2019.
36. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**:557–60.
37. Durham T, Guo J, Cowell W, Riley KW, Wang S, Tang D, et al. Prenatal PM2.5 exposure in relation to maternal and newborn telomere length at delivery. *Toxics* 2022;**10**.
38. Hautekiet P, Nawrot TS, Janssen BG, Martens DS, De Clercq EM, Dadvand P, et al. Child buccal telomere length and mitochondrial DNA content as biomolecular markers of ageing in association with air pollution. *Environ Int* 2021;**147**:106332.
39. Isaevska E, Fiano V, Asta F, Stafoggia M, Moirano G, Popovic M, et al. Prenatal exposure to PM10 and changes in DNA methylation and telomere length in cord blood. *Environ Res* 2022;**209**:112717.
40. Lee AG, Cowell W, Kannan S, Ganguri HB, Nentin F, Wilson A, et al. Prenatal particulate air pollution and newborn telomere length: effect modification by maternal antioxidant intakes and infant sex. *Environ Res* 2020;**187**:109707.
41. Mandakh Y, Oudin A, Erlandsson L, Isaxon C, Hansson SR, Broberg K, et al. Association of prenatal ambient air pollution exposure with placental mitochondrial DNA copy number, telomere length and preeclampsia. *Front Toxicol* 2021;**3**:659407.
42. Scholten R, Moller P, Jovanovic Andersen Z, Dehlendorff C, Khan J, Brandt J, et al. Telomere length in newborns is associated with exposure to low levels of air pollution during pregnancy. *Environ Int* 2021;**146**:106202.
43. Song L, Zhang B, Liu B, Wu M, Zhang L, Wang L, et al. Effects of maternal exposure to ambient air pollution on newborn telomere length. *Environ Int* 2019;**128**:254–60.
44. Carugno M, Borroni E, Fedrizzi L, Hoxha M, Vigna L, Consonni D, et al. Long- and short-term exposures to PM10 can shorten telomere length in individuals affected by overweight and obesity. *Life* 2021:11.
45. Dioni L, Hoxha M, Nordio F, Bonzini M, Tarantini L, Albetti B, et al. Effects of short-term exposure to inhalable particulate matter on telomere length, telomerase expression, and telomerase methylation in steel workers. *Environ Health Perspect* 2011;**119**:622–7.
46. Hou L, Wang S, Dou C, Zhang X, Yu Y, Zheng Y, et al. Air pollution exposure and telomere length in highly exposed subjects in Beijing, China: a repeated-measure study. *Environ Int* 2012;**48**:71–7.
47. Wong JY, De Vivo I, Lin X, Christiani DC. Cumulative PM(2.5) exposure and telomere length in workers exposed to welding fumes. *J Toxicol Environ Health, Part A* 2014;**77**:4410–55.
48. Isaevska E, Moccia C, Asta F, Cibella F, Gagliardi L, Ronfani L, et al. Exposure to ambient air pollution in the first 1000 days of life and alterations in the DNA methylol and telomere length in children: a systematic review. *Environ Res* 2021;**193**:110504.
49. Martens DS, Nawrot TS. Air pollution stress and the aging phenotype: the telomere connection. *Curr Environ Health Rep* 2016;**3**:258–69.
50. Petersen S, Saretzki G, von Zglinicki T. Preferential accumulation of single-stranded regions in telomeres of human fibroblasts. *Exp Cell Res* 1998;**239**:152–60.
51. Nieuwenhuijsen MJ, Khreis H, Triguero-Mas M, Gascon M, Dadvand P. Fifty shades of green: pathway to healthy urban living. *Epidemiology* 2017;**28**:63–71.
52. Factor-Litvak P, Susser E, Kezios K, McKeague I, Kark JD, Hoffman M, et al. Leukocyte telomere length in newborns: implications for the role of telomeres in human disease. *Pediatrics* 2016;**137**.
53. Okuda K, Bardeguez A, Gardner JP, Rodriguez P, Ganesh V, Kimura M, et al. Telomere length in the newborn. *Pediatr Res* 2002;**52**:377–81.
54. Frenck Jr RW, Blackburn EH, Shannon KM. The rate of telomere sequence loss in human leukocytes varies with age. *Proc Natl Acad Sci U S A* 1998;**95**:5607–10.
55. Liu L, Bailey SM, Okuka M, Munoz P, Li C, Zhou L, et al. Telomere lengthening early in development. *Nat Cell Biol* 2007;**9**:1436–41.
56. Schaetzlein S, Lucas-Hahn A, Lemme E, Kues WA, Dorsch M, Manns MP, et al. Telomere length is reset during early mammalian embryogenesis. *Proc Natl Acad Sci U S A* 2004;**101**:8034–8.

Public Health 215 (2023) 100–105

Contents lists available at ScienceDirect

Public Health

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

Original Research

Analyzing the impact of state gun laws on mass shootings in the United States from 2013 to 2021

Libby Tiderman [a], Nidele F. Dongmo [b], Katie Munteanu [c], Merlie Kirschenbaum [d], Lucy Kerns [e, *]

[a] Department of Mathematics, The University of Arizona, Tucson, AZ, USA
[b] Department of Economics and Business Management, Goucher College, Baltimore, MD, USA
[c] Department of Mathematics, Cerritos College, Norwalk, CA, USA
[d] Department of Mathematics, Brookdale Community College, Lincroft, NJ, USA
[e] Department of Mathematics & Statistics, Youngstown State University, Youngstown, OH, USA

ARTICLE INFO

Article history:
Received 9 August 2022
Received in revised form
19 October 2022
Accepted 1 December 2022
Available online 18 January 2023

Keywords:
Mass shootings
Gun regulation
Negative binomial generalized linear mixed model
Gun violence

ABSTRACT

Objectives: Mass shooting incidents have drastically increased in the United States in the last 10 years, with a disproportionate number of incidents occurring in some states. Gun laws vary greatly by state, but little research has been conducted to examine the association between the strength of state gun laws and mass shootings. This study aims to explore the aggregate effect of state gun laws on the rate of mass shooting incidents and fatalities.
Study design: This was a cross-sectional time series.
Methods: This study applied the negative binomial generalized linear mixed model to assess the impact of state gun laws restrictiveness—as measured by the total number of active gun laws—on the rate of mass shooting incidents and fatalities.
Results: The restrictiveness of state gun laws was significantly associated with the rate of mass shooting fatalities; specifically, for every 1 standard deviation (SD) increase in the state gun law restrictiveness score (i.e. for every additional 27 gun laws in place), the rate of mass shooting fatalities was decreased by 24% (P-value <0.0001), controlling for other predictor variables in the model. However, no significant association was found between the restrictiveness of state gun laws and rate of mass shooting incidents.
Conclusions: State gun laws may not decrease the number of mass shooting events, but they appear to help reduce the number of deaths when these mass shootings occur. Better data collection on mass shootings and further research on the impacts of specific gun laws are needed to help understand the effectiveness of gun laws and inform law-based interventions.

Introduction

The United States has substantially higher rates of gun-related civilian deaths compared with other major western countries.[1] Although mass shootings are relatively rare compared with other forms of gun violence, they have far-reaching impacts on people's mental health and sense of safety, whether they and their families were direct victims or witnesses of the incident. In the last 10 years, the number of mass shootings in the United States has drastically increased, with more than 30% of all mass shootings since 1966, having occurred after 2010.[2] The United States leads the world in this crisis, accounting for only 5% of the world's population, but 73% of global mass shootings occurred over the last two decades.[3] Mass shootings are violent acts stemming from many risk factors, ranging from individual-level factors (e.g. mental illness, lifestyle and family history) to population-level factors (e.g. gun regulation, media coverage, poverty percentage). Understanding the effects of relevant risk factors on mass shootings is an important step in developing successful intervention strategies for preventing such violence. Unfortunately, little is known regarding the associated risk factors of mass shootings and the cause of this rising trend in the United States because limited scientific research has been

* Corresponding author. Department of Mathematics and Statistics, Youngstown State University, One University Plaza, Youngstown, OH, 44555, USA. Tel.: +1 330 941 1412; fax: +1 330 941 3310.
E-mail address: xlu@ysu.edu (L. Kerns).

https://doi.org/10.1016/j.puhe.2022.12.001
0033-3506/

carried out to identify potential risk factors and examine their influences on mass shooting events.[4]

Gun regulation is one of the most important factors with the potential to reduce or prevent general gun-related violence or mass shootings in particular. Nevertheless, it is very controversial in the United States and has constantly been the subject of heated debates among policy makers. This is due in part to the fact that many of the possible effects of gun policies have only rarely—or never—been studied rigorously. There are limited federal gun laws in the United States, and states have the authority to enact their own gun laws regulating sales, purchase, possession, and storage of firearms. As a result, gun laws vary considerably across states, with some states being known to have many more gun restrictions than others. This, coupled with the seemingly disproportionate occurrence of mass shootings in some states, leads to an important question: "Is there an association between the occurrence of mass shootings and the strength of state-level gun laws?" Despite its importance, scientific research on examining the impacts of state gun laws on mass shootings has been relatively sparse and lacks consistency.[5]

Previous studies have examined the effects of some specific state-level gun laws on mass shootings with mixed results. In particular, state assault weapons bans appeared to have received more attention than other gun laws because assault weapons are uniquely lethal and are often the choice of public mass shooters. For example, Gius found that the ban on assault weapons had statistically significant and negative effects on mass shooting fatalities but not on mass shooting injuries.[6] However, a study by Webster et al. showed that this ban had little influence on fatal mass shootings.[7] Some studies have targeted different gun laws that may have impacts on mass shootings. For example, Duwe et al. focused on the Right-to-Carry laws and found no evidence that these laws are associated with mass shooting incidents and severity.[8] The study by Webster et al., which examined the association of mass shootings and a number of specific gun laws including the Right-to-Carry laws, also suggested that these laws are not associated with fatal mass shootings.[7] More studies have appeared in literature in addition to the ones mentioned here that investigated the impact of specific gun laws.[9–12]

Fewer studies have focused on the aggregate effect of all gun laws, rather than single laws. Most of these studies were primarily concerned with the association between firearm legislation and suicide[13–15] or the association between firearm legislation and homicide in general.[16,17] There has been particularly little research that specifically focused on mass shootings and the combined effect of all gun laws on this phenomenon, and the findings are controversial. Reeping et al. found that states with more permissive gun laws had higher rates of mass shooting incidents,[5] whereas both Duchesne et al.'s and Lin et al.'s studies showed that the strength of state gun laws did not affect rates of mass shooting incidents.[4,18] However, Duchesne et al. did find a significant correlation between gun law restrictiveness and the rate of mass shooting fatalities.[18]

Studies that focused on the cumulative impact of all gun laws often relied on some types of permissiveness-restrictiveness scale to quantify the strengths of gun laws in states. For example, Reeping et al.[5] used a rating between 0 (completely restrictive) and 100 (completely permissive) for the gun law strength, taken from the *Traveler's Guide to the Firearms Laws of the Fifty States*, whereas Duchesne et al.[18] obtained their law data from the Annual Gun Law Scorecard compiled by the Giffords Law Center, which assigned A–F letter grades based on the strengths of state gun laws, with A being the strongest gun laws and F being the weakest gun laws.[19] Because of the differences in the ways gun law permissiveness scales were created, there have been concerns about the reliability of these scales and the extent to which the choice of restrictiveness scale could affect the studies related to gun violence outcomes. Some efforts have been made to address these concerns; for example, Reeping et al. examined seven existing gun law restrictiveness scales—two rankings, two counts, and three scores—from different resources and found that these gun law restrictiveness scales are highly correlated, and the choice of a particular scale has little effect on study conclusions related to gun violence outcomes.[20]

In this study, we have gathered relevant information about mass shootings in the United States and the strength of state gun laws to explore trends in mass shootings and examine differences in gun law restrictiveness across states. We then sought to examine whether the restrictiveness of state gun laws was associated with mass shooting events.

Methods

Data sources

Mass shooting data from 2013 to 2021 were obtained from the Gun Violence Archive (GVA) database, which defines a mass shooting as an incident of gun violence in which four or more people are shot and/or killed, not including the perpetrator.[21] The GVA was established in 2013 as an independent and non-profit data collection organization to provide comprehensive data about gun violence. It collects and validates gun violence incidents from 7500 sources daily and has developed a database containing mass shootings in the United States from January 2013 through the present time using both automated queries and manual research.

To determine state gun law permissiveness, we used the data from the State Firearms Laws Database, which was created by Siegel et al. in an effort to provide researchers with the data necessary to evaluate the effectiveness of various firearm laws.[22] The database catalogs the presence or absence of each of 134 firearm safety laws in each state starting from 1991 and is constantly updated as states add and subtract gun laws. We collected the data for each year over an 8-year period (2013–2020, as 2021 is not available) for all states and counted the total number of active gun laws for each year in every state. These total law counts were then used as the state gun law permissiveness scales in this study. Owing to the omission of the District of Columbia in the State Firearms Laws Database, mass shootings within the District of Columbia were excluded from our research.

We included year as a covariate in all analyses to account for time-specific effects, and we also considered some socio-economic status variables that are often used in gun violence literature, including percent of population that is below the poverty level, percent of population that is Black, percent of population that has a 4-year college degree, median income, annual unemployment rate, percentage of population that is aged 15–34 years, and incarceration rate. All socio-economic data were gathered from the American Community Survey at the US Census Bureau,[23] except for the incarceration rate, which was obtained from the Bureau of Justice Statistics.[24] We further retrieved the state population data from the US Census Bureau to account for population differences among the states.[25]

Statistical methods

The annual mass shooting incidents and fatalities in the United States from 2013 to 2021 were calculated to assess the temporal trend of the mass shootings over this period. To understand how mass shooting incidents varied from state to state, we computed the population-based rates of mass shooting incidents and fatalities in each of the 50 states in all years. We constructed a boxplot to

display the distribution of the state-level gun law restrictiveness scores, represented by the average annual number of gun laws per state over the study period. We further constructed scatterplots between the state gun law restrictiveness scores and population-based rates of mass shooting incidents and fatalities.

We applied a generalized linear mixed model with a negative binomial distribution (negative binomial Generalized Linear Mixed Model (GLMM)) to investigate the impact of state-level gun law restrictiveness on mass shooting incidents and fatalities while accounting for sources of heterogeneity among states and dependence in counts of mass shooting incidents and fatalities within states. This model was chosen because both outcome variables (number of mass shooting incidents and fatalities) were found to be overdispersed count variables.

Specifically, we let Y_{it} represents the number of mass shooting incidents or fatalities at time t and state i, which is assumed to follow the negative binomial distribution:

$$Y_{it} \sim NB\left(\mu_{it}, \mu_{it} + \theta\mu_{it}^2\right) \quad (1)$$

where μ_{it} is the mean parameter and θ is the dispersion parameter that controls the amount of dispersion. The mean μ_{it} is related to the fixed state-level effects and random state effect via the following logarithm link function:

$$\log(\mu_{it}) = \log(P_{it}) + \boldsymbol{x}_{it}'\boldsymbol{\beta} + b_i \quad (2)$$

where $\log(P_{it})$ is an offset with P_{it} representing the state-specific population size, $\boldsymbol{\beta}$ is the 1×10 vector of fixed effects (including the intercept β_0) to represent the coefficients of the covariates $\boldsymbol{x}_{it}$, and b_i is the random state effect that is assumed to follow a normal distribution (i.e. $b_i \sim N(0, \sigma^2)$). In our model, the covariates vector $\boldsymbol{x}_{it}$ includes the gun law variable, year, and seven socio-economic covariates (i.e. percent of population that is below the poverty level, percent of population that is Black, percent of population that has a 4-year college degree, median income, annual unemployment rate, percentage of population that is aged 15–34 years, and incarceration rate). Note that year was included as a covariate to represent the underlying annual trend, and the gun law variable represents the average number of state gun laws per year, so it is a constant across years for each state. Furthermore, the state-specific random effect b_i was added to model the dependence in counts of mass shooting incidents and fatalities within each state because repeated counts were measured for the same state over time between 2013 and 2021.

In addition to the full models with all covariates, as specified in Equations (1) and (2), partially adjusted models were calculated by including the socio-economic covariates with coefficients in the full model that had P values less than 0.10. Note that the partially adjusted models included the variables year and gun law restrictiveness but limited inclusion of less influential socio-economic covariates.

All data analyses were performed in `R` software package, version 4.2.1 (Vienna University of Economics and Business).[26] To fit the negative binomial GLMM, we applied the `glmer.nb` function from the "`lme4`" package. We noticed that some of the covariates were measured on very different scales, which lead to convergence difficulties when fitting the model. Hence, we standardized the covariates before including them in the model to improve the performance of the optimization algorithm used in the `glmer.nb` function. The significance level was set as 5%.

Results

There were a total of 3585 mass shootings recorded between 2013 and 2021, with 3835 fatalities and 14,857 injuries, excluding mass shootings that occurred in Washington, DC. Louisiana had the highest rate of mass shooting incidents, with 3.46 per 100,000 residents, followed by Illinois (2.94 per 100,000 residents) and Mississippi (2.25 per 100,000 residents). Hawaii and North Dakota had no mass shooting incidents during the period.

As shown in Fig. 1, the number of mass shooting incidents has increased steadily over the study period, with the number having nearly tripled from 248 in 2013 to 676 in 2021. The largest 1-year increase occurred between 2019 (410 incidents) and 2020 (602 incidents)—a 31.9% increase in a single year. The number of mass shooting fatalities has followed a similar incline, rising from 275 in 2013 to 694 in 2021, and the largest single-year increase of 26.8% was recorded between 2020 (508 fatalities) and 2021 (694 fatalities).

On average, a state had 27.4 (SD = 26.8) state gun laws in place per year over the period from 2013 to 2020. California was found to have the most restrictive gun laws, with the average annual number of gun laws of 105.3, followed by Massachusetts (101.8) and Connecticut (86.3). The states with the most permissive gun laws were Mississippi (2.8), Idaho (2.9), and Alaska (3.0). Fig. 2 shows the boxplot of the distribution of the state-level gun law restrictiveness scores across all years. The average restrictiveness scores of state gun laws appeared to have a right-skewed distribution in which most of the states had an average score lower than the mean number of gun laws. Also, the aforementioned three states—California, Massachusetts, and Connecticut—were identified as outliers in the distribution with extremely large restrictiveness scores.

Fig. 3 displays scatterplots of gun law restrictiveness scores and rates of mass shooting incidents and fatalities. It can be seen that there was no apparent correlation between gun law restrictiveness scores and rates of mass shooting incidents (Fig. 3A); however, there appeared to be a negative correlation between gun law restrictiveness scores and rates of mass shooting fatalities (Fig. 3B), that is, states with more gun laws had lower rates of mass shooting fatalities.

The dispersion parameter θ in the negative binomial GLMM that examines the number of mass shooting incidents and fatalities was estimated to be 22.08 and 2.14, respectively, which indicates that both count data are overdispersed. Note that because the state population size was included as an offset term, the negative binomial GLMM models the rate of mass shooting incidents and fatalities (i.e. the number of incidents and fatalities per 100,000 residents). Also, the estimated coefficients for the predictors in the models were transformed into the rate ratio (RR) of mass shooting incidents and fatalities by an exponential function.

The results from the negative binomial GLMM with the number of mass shooting incidents as the outcome variable indicate that there was no statistically significant correlation between the rate of mass shooting incidents and state gun law restrictiveness (RR = 0.96, 95% confidence interval = 0.85, 1.09; P-value = 0.516), adjusted for the other predictors in the model. However, the negative binomial GLMM that examined the number of mass shooting fatalities shows that the state gun law restrictiveness had a statistically significant effect on the rate of mass shooting fatalities. Table 1 shows the results from the partially adjusted model with the number of mass shooting fatalities as the outcome variable.

The results in the table were calculated based on the standardized coefficients for the predictors in the partially adjusted models. The estimated RR of 0.76 for the gun law restrictiveness

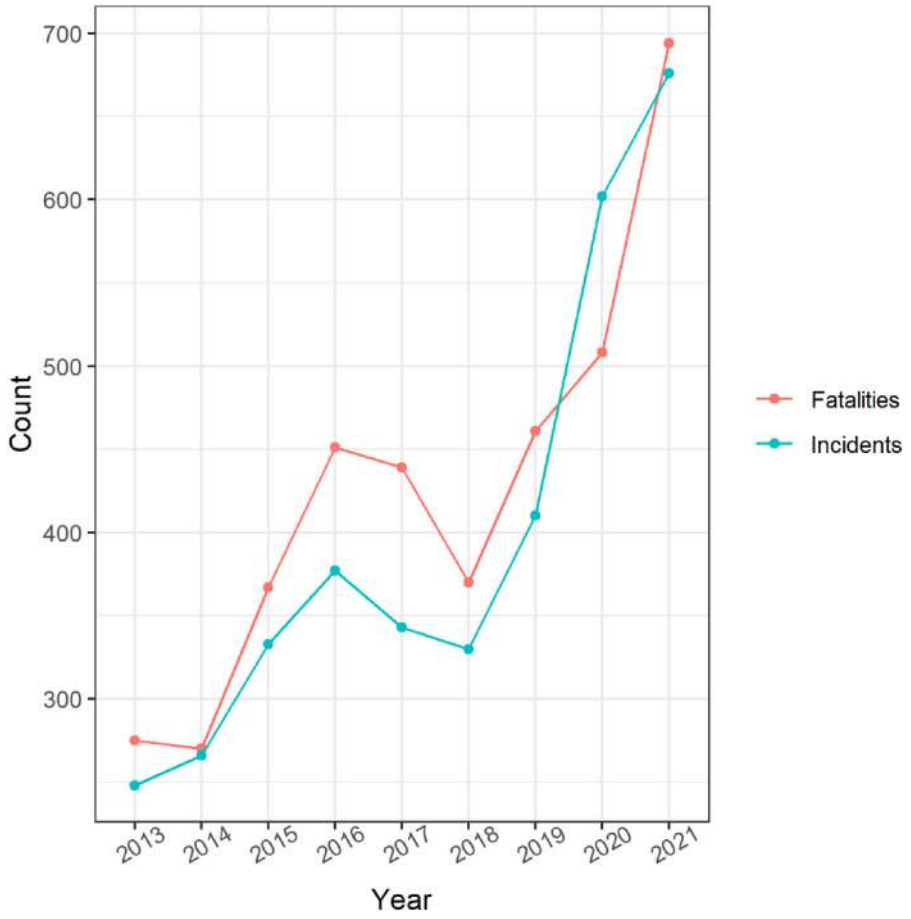


Fig. 1. Number of mass shooting incidents and fatalities in the United States for each year from 2013 through 2021.

score indicates that for every 1 SD increase in the state gun law restrictiveness score, the rate of mass shooting fatalities is decreased by 24% (RR = 0.76, 95% confidence interval = 0.66, 0.87; *P*-value <0.0001), controlling for other predictor variables in the model. In other words, having additional 27 gun laws (the SD of gun law scores is 26.8) in place per year is associated with a 24% decrease in the rate of mass shooting fatalities. In addition, the estimates of RR for the predictor, year, indicate a statistically significant time trend (RR = 1.26, 95% confidence interval = 1.14, 1.39; *P*-value <0.0001).

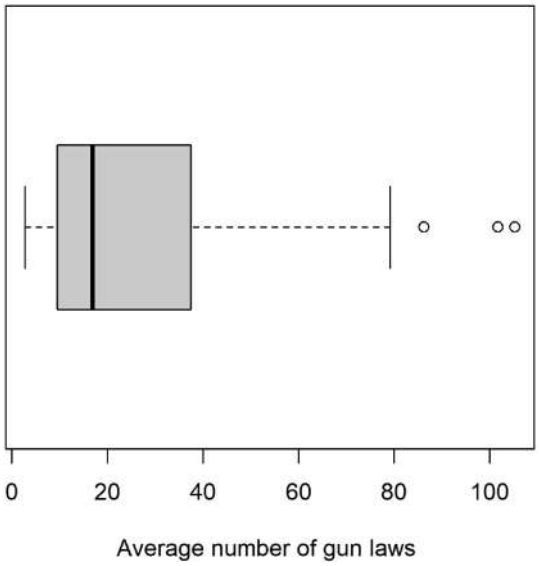


Fig. 2. Boxplot of average number of gun laws per state across all years.

Discussion

Our study found that more restrictive gun legislation is significantly associated with lower rates of fatal mass shootings. This result is congruent with other literature on this topic.[6,11,18] In addition, our findings that the restrictiveness of state firearm legislation is not significantly correlated with the rate of mass shooting incidents is consistent with the work of Duchesne et al.[18]; however, these findings conflict with the results of Reeping et al.[5] and Siegel et al.[12]

One possible explanation for the discrepancy in the findings of our research and other studies on mass shootings could be the difference in the definitions used for a mass shooting. The GVA defines a mass shooting as an event in which four or more are shot, but not necessarily killed, not including the shooter, whereas most prior studies relied on data from other sources using different definitions; for example, the Federal Bureau of Investigation defines a mass shooting as a single event in which four or more people are killed, excluding the shooter. The lack of consistent definitions makes it infeasible to effectively compare the research results of different studies on mass shootings, resulting in limited

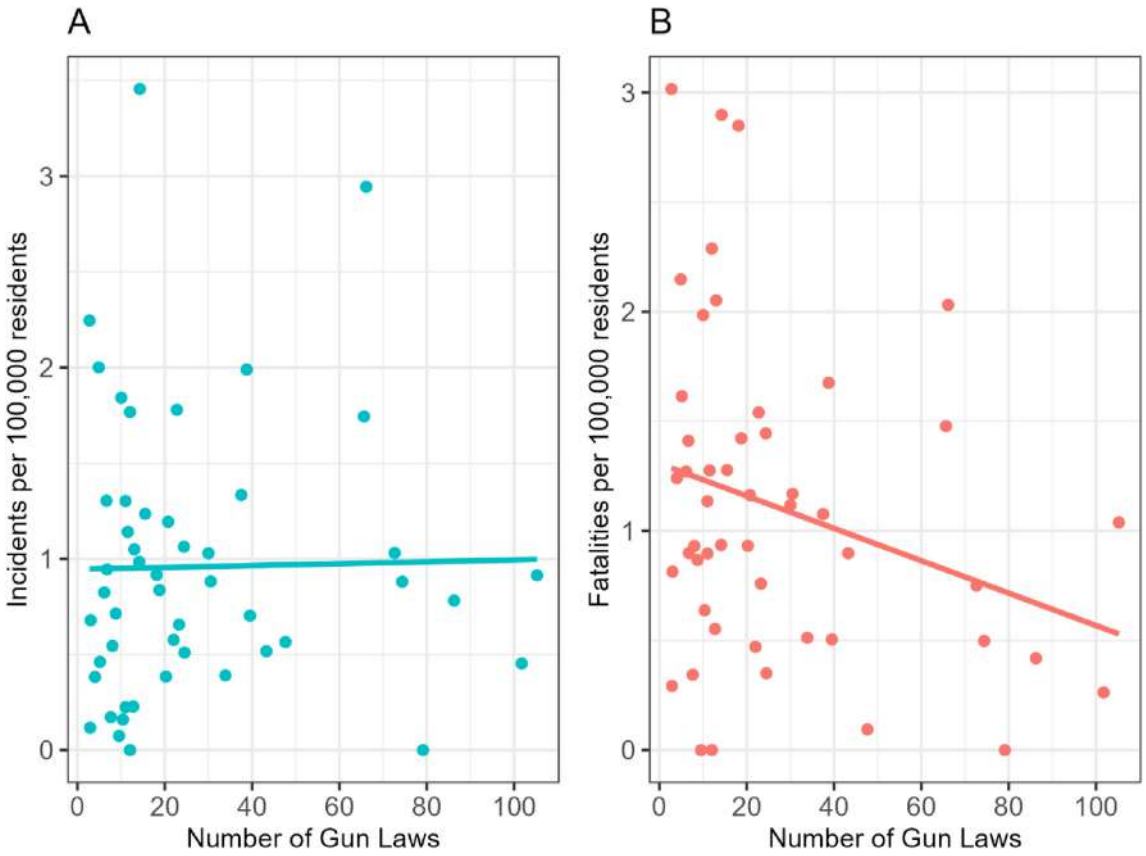


Fig. 3. Scatterplots of the relations between gun law restrictiveness scores and rates of mass shooting incidents (A) and fatalities (B).

Table 1
Estimates of rate ratio for the mass shooting fatalities.

Variable	Rate ratio	95% CI
Gun law restrictiveness score	0.76	0.66, 0.87
Year	1.26	1.14, 1.39
Unemployment rate	1.36	1.16, 1.60
Percent Black	1.23	1.06, 1.43

CI, confidence interval.

understanding of which measures should be taken to address this public health crisis. Therefore, it is highly recommended to have a standard definition for a mass shooting to help ensure the comparability of research findings related to mass shootings and a better understanding of this crisis.

We measured the gun law restrictiveness by counting the total number of gun laws in a state. An important assumption of such a strategy is that each gun law has an equal impact on mass shootings. One could argue that every individual gun law may have different impacts, and hence, a different weight should be assigned to represent their respective strengths. In fact, some weighted scoring systems have been proposed by some organizations (e.g. the Brady Center[27]). However, these scoring systems have their own disadvantages, and none of them have been validated.[17] This might be partly attributable to the fact that limited research has been done targeting specific gun legislation, and moreover, the findings from existing studies focusing specifically on individual laws are not always consistent, as pointed out earlier. Therefore, more studies are needed to understand how specific gun laws affect the mass shootings and develop more reliable measures to quantify the restrictiveness or permissiveness of state gun control laws.

Although our study has found a significant association between the strength of state gun laws and mass shooting fatalities, this association could have been confounded by some unaccounted factors. One such factor could be gun ownership because one way that gun legislation could help reduce gun-related violence is through reducing access to guns. Previous work has shown that levels of gun ownership were significantly associated with mass shooting rates.[5] Another important factor that may influence the impact of gun legislation on mass shootings is the enforcement of gun laws. Unfortunately, it is difficult to study and measure the level of enforcement, such as whether and how a gun law is enforced, and/or if there is possible exploitation of legal loopholes. This is mainly because of insufficient data or lack of adequate statistical methods. In addition, although we considered a number of state-level factors in the present study, the relationship between gun law restrictiveness and mass shootings might be confounded by other factors that are not considered but may be relevant (e.g. percentage of female-headed households, percent of high school graduation).

Limitations

There are a number of limitations of this study. First of all, we used a simplified approach of counting the number of state-level gun laws to calculate gun law restrictiveness scores. This scoring method, which assumes an equal impact of each gun law on mass shootings, has not been validated. Second, it has been argued that mass shootings resulting from more conventional crimes—such as domestic violence, armed robbery, and gang-related crimes—are contextually distinct from indiscriminate incidents in which victims were selected randomly in relatively public places, and as

such, should be treated separately. The GVA database contains a broad range of mass shooting incidents, and the task of identifying them as specific types of events (e.g. domestic or indiscriminate) to a reasonable certainty is challenging because there is insufficient information available about the perpetrator or his motives for some of the reported incidents. Therefore, we were unable to divide mass shooting events into distinct groups and analyze them separately. These separate analyses, in addition to the overall analysis conducted in the present study, could have shed more light on the impacts of gun law strengths on different types of mass shootings and their associated policy implications. Finally, although our study has found a significant association between the strength of state gun laws and mass shooting fatalities, we cannot determine a cause-and-effect relationship. Further research is needed to have a better understanding of the nature of this association.

Conclusions

The restrictiveness of state gun laws—as measured by the total number of laws—is significantly associated with the rate of mass shooting fatalities but not with the rate of mass shooting incidents in the United States. Better data collection on mass shootings—such as implementing a standardized definition of mass shootings—and further research on the impacts of specific gun laws are needed to help understand the effectiveness of gun laws and inform law-based interventions.

Author statements

Ethical approval

Not applicable because the data used in the study are freely available to the public.

Funding

This work was supported by the National Science Foundation (NSF Grant DMS-2050789). It was completed during the Youngstown State University Beginning Undergraduates Mathematical Research Preparation (YSU-BUMP) Summer REU program.

Competing interests

None declared.

References

1. Pritchard C, Parish M, Williams RJ. International comparison of civilian violent deaths: a public health approach to reduce gun-related deaths in US youth. *Public Health* 2020 Mar;**180**:109–13.
2. National Institute of Justice. *Public Mass Shootings: Database Amasses Details of a Half Century of U.S. Mass Shootings with Firearms, Generating Psychosocial Histories* [Internet]. 2022 [cited 2022 Jun 28]. Available from: https://nij.ojp.gov/topics/articles/public-mass-shootings-database-amasses-details-half-century-us-mass-shootings.
3. Silva JR. Global mass shootings: comparing the United States against developed and developing countries. *Int J Comp Appl Crim Justice* 2022 Mar 21:1–24.
4. Lin PI, Fei L, Barzman D, Hossain M. What have we learned from the time trend of mass shootings in the U.S. *PLoS One* 2018 Oct 18;**13**(10):e0204722.
5. Reeping PM, Cerdá M, Kalesan B, Wiebe DJ, Galea S, Branas CC. State gun laws, gun ownership, and mass shootings in the US: cross sectional time series. *BMJ* 2019 Mar 6:l542.
6. Gius M. The impact of state and federal assault weapons bans on public mass shootings. *Appl Econ Lett* 2015 Mar 4;**22**(4):281–4.
7. Webster DW, McCourt AD, Crifasi CK, Booty MD, Stuart EA. Evidence concerning the regulation of firearms design, sale, and carrying on fatal mass shootings in the United States. *Criminol Public Policy* 2020 Feb 30;**19**(1):171–212.
8. Duwe G, Kovandzic T, Moody CE. The impact of right-to-carry concealed firearm laws on mass public shootings. *Homicide Stud* 2002 Nov 24;**6**(4):271–96.
9. Nagin DS, Koper CS, Lum C. Policy recommendations for countering mass shootings in the United States. *Criminol Public Policy* 2020 Feb 10;**19**(1):9–15.
10. Koper CS, Johnson WD, Nichols JL, Ayers A, Mullins N. Criminal use of assault weapons and high-capacity semiautomatic firearms: an updated examination of local and national sources. *J Urban Health* 2018 Jun 2;**95**(3):313–21.
11. Klarevas L, Conner A, Hemenway D. The effect of large-capacity magazine bans on high-fatality mass shootings, 1990–2017. *Am J Public Health* 2019 Dec;**109**(12):1754–61.
12. Siegel M, Goder-Reiser M, Duwe G, Rocque M, Fox JA, Fridel EE. The relation between state gun laws and the incidence and severity of mass public shootings in the United States, 1976–2018. *Law Hum Behav* 2020 Oct;**44**(5):347–60.
13. Conner KR, Zhong Y. State firearm laws and rates of suicide in men and women. *Am J Prev Med* 2003 Nov;**25**(4):320–4.
14. Cantor CH, Slater PJ. The impact of firearm control legislation on suicide in Queensland: preliminary findings. *Med J Aust* 1995 Jun;**162**(11):583–5.
15. Rodríguez Andrés A, Hempstead K. Gun control and suicide: the impact of state firearm regulations in the United States, 1995–2004. *Health Policy* 2011 Jun;**101**(1):95–103.
16. Price JH, Thompson AJ, Dake JA. Factors associated with state variations in homicide, suicide, and unintentional firearm deaths. *J Community Health* 2004 Aug;**29**(4):271–83.
17. Fleegler EW, Lee LK, Monuteaux MC, Hemenway D, Mannix R. Firearm legislation and firearm-related fatalities in the United States. *JAMA Intern Med* 2013 May 13;**173**(9):732.
18. Duchesne J, Taghavi S, Toraih E, Simpson JT, Tatum D. State gun law grades and impact on mass shooting event incidence: an 8-year analysis. *J Am Coll Surg* 2022 Apr 18;**234**(4):645–51.
19. *Annual Gun Law Scorecard [Internet]* [cited 2022 Jul 26]. Available from: https://giffords.org/lawcenter/resources/scorecard.
20. Reeping PM, Morrison CN, Rudolph KE, Goyal MK, Branas CC. A comparison and analysis of seven gun law permissiveness scales. *Inj Epidemiol* 2021 Dec 18;**8**(1):2.
21. *Gun Violence Archive [Internet]* [cited 2022 May 26]. Available from: www.gunviolencearchive.org.
22. *State Firearm Law Database [Internet]*. Available from: www.statefirearmlaws.org/state-state-firearm-law-data.
23. *American Community Survey. United States Census Bureau, Washington, DC* [Internet]. [cited 2022 Jul 31]. Available from: https://www.census.gov/programs-surveys/acs.
24. *Carson AEP in 2020. Carson [Internet]* [cited 2022 Jul 26]. Available from: https://bjs.ojp.gov/content/pub/pdf/p20st.pdf.
25. *US Census Bureau [Internet]*. Available from: https://www.census.gov/programs-surveys/decennial-census/decade/2020/2020-census-results.html.
26. R Core Team. *R: a language and environment for statistical computing* [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available from: https://www.r-project.org/.
27. *Brady Campaign to Prevent Gun Violence. Legislative scorecards [Internet]*. Available from: http://www.bradycampaign.org/stategunlaws/.

Public Health 215 (2023) 56–65

Contents lists available at ScienceDirect

Public Health

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

Original Research

Challenges in moving toward universal health coverage: rising cost of outpatient care among Vietnam's insured rural residents, 2006–2018

A. Sepehri [a, *], K.N. Minh [b], P.H. Vu [c]

[a] Department of Economics, University of Manitoba, Winnipeg, Canada
[b] Department of Economics and Management, Thang Long University, Hanoi, Viet Nam
[c] School of Banking & Finance, National Economics University, Hanoi, Viet Nam

ARTICLE INFO

Article history:
Received 6 July 2022
Received in revised form
30 November 2022
Accepted 1 December 2022
Available online 13 January 2023

Keywords:
Universal health coverage
Out-of-pocket expenditure
Outpatient care
Health insurance
Vietnam

ABSTRACT

Objectives: The aim of this study was to assess temporal trends in out-of-pocket (OOP) expenditures per outpatient contact by the insured residents in rural Vietnam.
Study design: This was a repeated cross-sectional study.
Methods: Seven biennial waves from the Vietnam's Household Living Standard Survey covering the period 2006–2018 and a two-part model were used to assess temporal trends in OOP expenditures and its variations across various health facilities while controlling for a wide array of individual- and household-specific characteristics.
Results: The pattern of health facility utilization shifted steadily from commune health centers toward higher level government hospitals and private health facilities between 2006 and 2018. The regression results indicated an upward trend in the amount of OOP expenditures, with the amount of OOP expenditures incurred per outpatient contact being 40.3% higher in 2010–2012 than in 2006–2008 and by as much as 155.5% higher in 2018. These high-cost pressures were attenuated by 63%–65% when accounting for the types of health facility contacted. The cost inflation was more pronounced for care sought at higher level government hospitals and private hospitals than at other health facilities.
Conclusion: The cost of accessing outpatient care rose sharply between 2006 and 2018, particularly for visits involving higher level government hospitals and private hospitals. These findings suggest that beside expanding the coverage over the transition path to universal coverage, efforts should be directed at reforming Vietnam's hospital-centric and fragmented delivery system as a way of containing costs and ensuring financial sustainability of social health insurance system.

Introduction

In recent years, a growing number of low- and middle-income countries (LMICs) have implemented national health insurance (SHI) schemes that aim at advancing the transition to universal health coverage (UHC), defined as securing "access to key promotive, preventive, curative and rehabilitative health interventions for all at an affordable cost".[1] The three core dimensions of UHC are population coverage (who is covered), service coverage (which benefits are covered), and financial coverage (what proportion of health service expenses are covered).

Recent experience from LMICs shows that although many countries have been able to increase substantially population coverage and access to healthcare services, the overall evidence on financial risk protection and managing cost pressures is less clear.[2–7] The impact of health insurance on OOP expenditures is found to be modest at best, especially for many illnesses, which are of a less catastrophic nature and require ambulatory care.[8–13] Many countries continue to grapple with the challenges of controlling cost pressures without eroding coverage and undermining further financial protection.[8,14]

The risk of cost escalation is particularly high in settings where the delivery system remains hospital-centered and fragmented, where public hospitals enjoy some degree of autonomy to manage their financial resources, and where inefficiencies in resource allocations pertaining to the scope of benefits package, provider payment mechanisms, and purchasing continue to pose a threat to

* Corresponding author. Department of Economics, University of Manitoba, Fort Garry Campus, Winnipeg, MB, R3T 5V5, Canada.
E-mail addresses: Ardeshir.Sepehri@umanitoba.ca (A. Sepehri), khacminh@gmail.com (K.N. Minh), phuongvh@neu.edu.vn (P.H. Vu).

https://doi.org/10.1016/j.puhe.2022.12.002
0033-3506/

cost containment and the financial sustainability of social health system.[14–19] In this hospital-centric system, the quality of routine primary health care is often far from adequate, with the primary care facilities failing to meet the primary care needs of the population and their rising expectations, particularly in areas of screening and management of non-communicable diseases and concomitant increases in care complexity.[19–24] A weak primary care sector compounds the risk of a cost escalation as patients with adequate resources bypass primary care facilities and seek care at already over-crowded higher level hospitals.[15,19,25–30] Patients bypassing the local primary health center and seeking care at private facilities reported spending 2.3 times more and as much as 10.3 times more when seeking care at a hospital.[28] Obtaining care from higher level hospitals may become even more costly for rural residents who often incur longer travel times and waiting times than their urban counterparts.[31,32]

Although there is a large and growing literature on the impact of SHI schemes on the utilization of healthcare services and financial protection,[3,4,12] much of this literature is cross-sectional in nature, focusing on financial protection provided by health insurance schemes in a short time frame, typically 1 year. Comparatively, little is known about the trend in cost of accessing healthcare services over the transition path to UHC.[10,33–36] The primary aim of this article was to fill this gap by using Vietnam as a case study and investigate trends in OOP expenditures per outpatient contact among the insured rural residents. More specifically, this article attempts to answer three questions: (1) What was the extent of cost pressure in OOP expenditures incurred per outpatient contact by the insured over the study period? (2) How much of this increase in cost inflation was mediated by changes in the pattern of outpatient care utilization? (3) Did the cost inflation vary across health facilities? This was achieved using seven biennial waves from the Vietnam's Household Living Standard Survey (VHLSS) covering the period 2006–2018 and a two-part model (TPM). As health expenditures data for those with any healthcare use typically exhibit a substantial mass at zero, a TPM allows for the possibility that zeros and positive values are generated by different mechanisms and investigates separately the influence of covariates on the likelihood of zero vs non-zero OOP expenditures and the amount of OOP expenditure. It is important from a policy perspective to evaluate the trend in OOP expenditure over the transition path to UHC, particularly considering that many health insurance schemes in LMICs, including Vietnam, are continually evolving as are the broader health systems context. To our knowledge, the present article is the first study to investigate temporal trends in OOP expenditures per outpatient contact by the insured residents in rural Vietnam.

Vietnam has made impressive progress toward UHC over the past three decades through a series of incremental reforms to its SHI system that expanded population coverage from 5% in its year of inception (1993) to 60% in 2010 and nearly 87% in 2018.[37,38] The government roadmap toward UHC aims at achieving more than 90% coverage by 2020 and 95% by 2025.[38] Crucial to the rapid increase in the coverage rate was a government commitment to the goal of UHC and subsequent large increases in government budgetary transfers that were used to fully subsidize insurance premiums for the poor, ethnic minorities in disadvantaged regions, social assistance beneficiaries, and children under six and to partially subsidize the premiums for the near-poor school children and students as well as the rest of the informal sector.[17,37] By 2018, about 36%–40% of SHI members were fully subsidized and 30% partially subsidized, and these large fiscal transfers accounted for 25% of the health insurance fund, up from 5% in 2005.[37] A regulatory framework was gradually developed, and the first Health Insurance Law came into force in 2009, merging various insurance schemes and making SHI the primary mechanism for achieving UHC.[17] The benefit package covers most outpatient and inpatient contacts, laboratory examinations, diagnostic tests, and drugs listed as reimbursable by the ministry of health. The insured benefits can mainly be accessed at the assigned public health facilities and a small but growing number of private health facilities with contracts with the national insurer, Vietnam Social Security (VSS). The co-insurance rate for covered services is 20% for most insured groups and much higher for those bypassing their designated lower level public hospitals for a higher level health facility.

However, Vietnam continues to face many challenges associated with reducing OOP spending and reforming its costly and inefficient hospital-centric delivery structure.[38–41] Despite the expansion of health insurance coverage and the generous outpatient and inpatient benefits package provided by SHI, OOP spending by households remains persistently high at just under 40%.[38]

Methods

Data

The data used in this study were from the VHLSS for the years 2006, 2008, 2010, 2012, 2014, 2016, and 2018. The VHLSS is a large, nationally representative survey conducted by the General Statistical Office of Vietnam every 2 years. The survey targets the civilian, non-institutionalized population of Vietnam at the household level and uses a three-stage (communes/wards, enumeration areas, households) stratified random cluster sampling design to collect information through face-to-face interviews with household heads and key commune officials in 3063 (3313) communes/wards in 2006–2010 (2012–2018). The samples are representatives of the whole country, eight regions, urban/rural areas, and provinces. The numbers of households included in each sample were 9189 in 2006–2008 and 9399 in 2010–2018. The General Statistical Office also provides sample weights accounting for the survey design and for non-response. Using a 12-month recall period the health section of each survey collects information on household health seeking behavior, the types of health facility contacted, OOP expenditures, reasons for seeking care, health insurance status, and whether they used their insurance card when seeking care at a facility. The sample for this study was restricted to the insured rural residents who accessed their insurance benefits when seeking outpatient care. Once we match individual records of outpatient visits (39,781) with other variables, our sample consists of 39,762 outpatient visits over the period 2006–2018.

Study variables

The dependent variable was real OOP expenditures per outpatient visit, including consultation, diagnosis, medication, travel, and accommodation expenses incurred by the insured patients for 12 months. Nominal OOP expenditures were deflated using the consumer price index (2010 = 100). Consistent with the previous literature, the main covariates of interest included a wide range of individual- and household-level factors, including the types of outpatient visit, the types of health facility visited, age, marital status, education, household economic status, ethnicity, and geographical location. Health facilities were categorized into five groups: commune health centers (CHCs), district hospitals (reference category), the higher level public hospitals, private hospitals, and private clinics. As the number of outpatient contacts at regional general polyclinics was small, we included regional general polyclinics under district hospitals. The types of outpatient visits were represented by four dummies: vaccination, family planning, check-ups (reference category), and medical treatment. Age was presented by a continuous

variable; gender by a dummy variable that takes the value of one when the individual is male; and marital status by a dummy variable that equals one if the individual is currently married. The level of education of the respondent was measured by a set of the highest educational attainment dummies: below primary (reference category), primary, secondary, high school, and postsecondary. The education variable for children aged ≤15 years refers to education of the head of the household. Household economic status was measured by the real per capita consumption expenditure quintile. As household consumption expenditure comprises both monetary and in-kind payments on all goods and services, and the money value of home-made products, it is conventionally viewed to be the preferred measure of living standards in developing economies.[46] Ethnicity was measured by a dummy variable that takes the value of one if an individual is a member of the ethnic majority *kinh* or Chinese ethnic community and geographical location by eight regional dummies; Red River Delta (reference category), Northeast, Northwest, North Central Coast, South Central Coast, Central Highlands, Southeast, Mekong Delta. A set of year dummies was used to account for year-specific changes: 2006–2008 (reference category), 2010, 2012, 2014, 2016, and 2018. Finally, to reduce the distorted effect of the upper outlying OOP expenditures on the results, a dummy variable (Dum_outliers) was included, taking the value of one if ln(OOP expenditures) > (Q3 + 1.5 × interquartile range), where Q3 is the 75th percentile (upper quartile), Q1 is the 25th percentile (lower quartile), and the interquartile range is Q3–Q1. These outliers accounted for 2.6% of the sample. The definition of these variables and their summary statistics are provided in Table 1.

Table 1
Definitions and summary statistics (unweighted).

Variable	Description	Total sample (N = 39,762)	Incurring positive OOP expenditures (N = 29,318)
		Mean (Std. dev.)	Mean (Std. dev.)
Dependent variables			
Incurring positive OOP expenditures	=1 if incurred positive OOP expenditures, 0 otherwise	0.737 (0.440)	
Log of real out-of-pocket expenditure per outpatient contact	ln(real OOP expenditures per outpatient contact for positive expenditures ['000 2010 VND])		4.2494 (1.482)
Control variables			
Types of health facility			
Commune health center	=1 if contacted a commune health center, 0 otherwise	0.403 (0.490)	0.289 (0.453)
District hospital (ref.)	=1 if contacted a district hospital, 0 otherwise	0.324 (0.468)	0.359 (0.468)
Higher level govt. hospital	=1 if contacted a higher level government hospital, 0 otherwise	0.148 (0.355)	0.187 (0.390)
Private hospital	=1 if contacted a private hospital, 0 otherwise	0.024 (0.152)	0.031 (0.173)
Private clinic	=1 if contacted a private clinic, 0 otherwise	0.102 (0.302)	0.134 (0.341)
Types of outpatient visit			
Check-ups (ref.)	=1 if regular check-ups, 0 otherwise	0.243 (0.429)	0.214 (0.410)
Vaccination	=1 if vaccination, 0 otherwise	0.054 (0.225)	0.025 (0.157)
Medical treatment	=1 if medical treatment, 0 otherwise	0.682 (0.466)	0.738 (0.440)
Prenatal/family planning	=1 if prenatal/family planning, 0 otherwise	0.022 (0.146)	0.023 (0.150)
Age	years	37.390 (25.518)	39.388 (24.957)
Age-squared	years	2049.203 (2047.813)	2174.237 (2036.098)
Sex	=1 if male, 0 if female	0.428 (0.495)	0.424 (0.494)
Marital status	=1 if married, 0 otherwise	0.471 (0.499)	0.491 (0.500)
Level of educational attainment			
Below primary (ref.)	=1 if highest education is some primary or none, 0 otherwise	0.341 (0.474)	0.324 (0.468)
Primary	=1 if highest education is primary, 0 otherwise	0.246 (0.431)	0.253 (0.435)
Secondary	=1 if highest education is secondary, 0 otherwise	0.265 (0.441)	0.271 (0.444)
High school	=1 if highest education is high school, 0 otherwise	0.110 (0.312)	0.111 (0.315)
Postsecondary	=1 if highest education is college or above, 0 otherwise	0.038 (0.191)	0.041 (0.199)
Household consumption expenditure quintiles			
Quintile 1 (the poorest) (ref.)	=1 if expenditure quintile is 1, 0 otherwise	0.293 (0.455)	0.242 (0.428)
Quintile 2	=1 if expenditure quintile is 2, 0 otherwise	0.225 (0.417)	0.228 (0.419)
Quintile 3	=1 if expenditure quintile is 3, 0 otherwise	0.203 (0.402)	0.218 (0.413)
Quintile 4	=1 if expenditure quintile is 4, 0 otherwise	0.172 (0.377)	0.190 (0.393)
Quintile 5	=1 if expenditure quintile is 5, 0 otherwise	0.107 (0.310)	0.122 (0.327)
Regions			
Red River Delta (ref.)	=1 if resides in this region, 0 otherwise	0.151 (0.358)	0.161 (0.368)
Northeast	=1 if resides in this region, 0 otherwise	0.151 (0.358)	0.127 (0.332)
Northwest	=1 if resides in this region, 0 otherwise	0.062 (0.241)	0.043 (0.203)
North Central Coast	=1 if resides in this region, 0 otherwise	0.104 (0.305)	0.101 (0.301)
South Central Coast	=1 if resides in this region, 0 otherwise	0.091 (0.287)	0.087 (0.281)
Central Highlands	=1 if resides in this region, 0 otherwise	0.082 (0.274)	0.077 (0.267)
Southeast	=1 if resides in this region, 0 otherwise	0.094 (0.292)	0.102 (0.303)
Mekong Delta	=1 if resides in this region, 0 otherwise	0.265 (0.441)	0.302 (0.459)
Year dummies			
2006–2008 (ref.)	=1 if year is 2006 or 2008, 0 otherwise	0.207 (0.405)	0.207 (0.405)
2010	=1 if year is 2010, 0 otherwise	0.145 (0.352)	0.138 (0.344)
2012	=1 if year is 2012, 0 otherwise	0.147 (0.354)	0.147 (0.354)
2014	=1 if year is 2014, 0 otherwise	0.152 (0.359)	0.152 (0.359)
2016	=1 if year is 2016, 0 otherwise	0.169 (0.375)	0.169 (0.375)
2018	=1 if year is 2018, 0 otherwise	0.180 (0.384)	0.187 (0.390)
Outliers	=1 if ln(OOP expenditures) >(quartile 3 + 1.5 × interquartile range) in each of the survey rounds		0.026 (0.158)

OOP, out-of-pocket.

Statistical analysis

Healthcare expenditure data, for those with any healthcare use, are challenging to model because their distributions typically display substantial skewness to the right and a large fraction of observations with zero expenditures.[42] A TPM is commonly used to address challenges posed by healthcare expenditure data.[12,42–45] By separating decision-making into two independent steps, the TPM allows for the possibility that zeros and positive values are generated by different mechanisms. We used a binary logistic regression to analyze the probability of incurring any OOP expenditures, and a generalized linear model (GLM) with a log link and gamma distribution to analyze the amount of OOP expenditures. The selection of link function and distribution family was guided by the standardized specification tests, namely, a Box–Cox test and a modified Park test.[42]

More specifically, using a logistic regression, the probability of having positive OOP expenditures for individual i using facility j ($OOPE_{ij}$) is specified as a function of the types of health facility contacted ($facility_j$), a set of observed individual- and household-specific control variables (x_{ij}), and a set of year dummies (dum_year_t):

$$\text{logit}(OOPE_{ij}) = \alpha_0 + \Sigma_j\, \alpha_j\, facility_{ij} + \Sigma_k\, \alpha_k x_{ik} + \Sigma_t\, \alpha_t\, dum\_year_t + \varepsilon_{ij} \quad (1)$$

where a_0 is the constant or intercept and ε_{ij} the error term. The same explanatory variables are used in the second part of the model, estimating the amount of OOP expenditures for individuals reporting positive level of OOP expenditures:

$$\ln(OOPE_{ij} \mid OOPE_{ij} > 0) = \beta_0 + \Sigma_j\, \beta_j\, facility_{ij} + \Sigma_k\, \beta_k x_{ik} + \Sigma_t\, \beta_t\, dum\_year_t + \nu_{ij} \quad (2)$$

To assess the extent to which the time trend gradients in the likelihood of having positive OOP expenditures and its amount, as measured by α_t and β_t, respectively, were mediated by the observed changes in the pattern of health facility utilization, equations (1) and (2) were estimated with and without controlling for the types of health facility. Regression equation (2) was also extended in two important ways. First, to access the extent to which the cost inflation varied among health facilities over the study period, types of health facility variables in regression equation (2) were replaced by a set of two-way interaction terms involving types of health facility and each year dummy variable. Second, to assess the differential influence of household expenditure quintiles on the amount of OOP expenditures incurred across various health facilities, household expenditure quintiles in regression equation (2) were replaced by a set of interaction terms between each type of health facility and household expenditure quintile.

To assess the robustness of the results, we also applied a two-level (individual and commune/ward) random intercept logistic regression model to analyze the probability of incurring any OOP expenditures. To the extent that the likelihood of incurring positive OOP expenditure is influenced by the unobserved characteristics of the community, the likelihood of an insured member incurring any OOP expenditures is likely to be correlated among the community members. In this case, the application of standard binary regression models such as a logistic model could lead to bias.[47] We estimated a two-level random intercept logistic regression model without including the observed covariates and calculate the intracommune correlation (ρ). The estimated ρ was very low (0.013), indicating no significant degree of homogeneity in the likelihood of having positive OOP expenditures among the observed responses within a commune/ward.

As the VHLSS data uses a three-stage, stratified cluster sampling methodology, the standard errors of the estimated coefficients were corrected for potential intracommune correlation resulting from the clustering of responses by the primary sampling unit (commune). We also applied sampling weights to produce unbiased population estimates. Stata version 14.1 (StataCorp, College Station, TX) was used for all data analysis.

Results

The real OOP expenditures by those incurring positive OOP expenditure rose by almost 2.8-folds between 2006 and 2018, and the cost inflation was more pronounced for contacts at higher level government hospitals than at other health facilities (Fig. 1). Given the small numbers of outpatient contacts at private hospitals in 2006–2008 (18), 2010–2012 (84), OOP per outpatient contacts at these facilities should be interpreted with caution.

The overall growth rate of OOP expenditures reported in Fig. 1 reflects changes in the pattern of outpatient care utilization over the period under consideration as well as provider-specific cost inflation. Over the study period, the share of outpatient contacts at CHCs dropped sharply from over half of all outpatient contacts in 2006–2008 to 28.6% in 2018, whereas the share of contacts at higher level government hospitals rose from 11.3% to 19.8% and private clinics from 1.5% to 13.1% (Fig. 2). The shift in the pattern of outpatient care utilization was more pronounced for the non-poor than for the poor. CHCs remained the dominant source of care for the poor patients, accounting for half of all outpatient contacts by the poor in 2018 vs 70.4% in 2006–2008.

The results from estimating the TPM with and without controlling for the types of health facilities contacted are reported in Table 2. Coefficient estimates from the first part (logit) indicate the likelihood of having positive OOP expenditures and those of the second part (GLM) the level of OOP expenditures. To facilitate interpretation, the estimated coefficients of the logistic model are converted into odds ratios. In the semilogarithmic GLM regression equation of Table 2 where the outcome variable is the natural logarithm of OOP expenditure per outpatient contact, the influence of a covariate on OOP expenditures is given by ($e^{\beta} - 1$). The results for year dummies indicate a far steeper time trend gradient in the amount of OOP expenditures than in the odds of having any OOP expenditures, with the amount of OOP expenditures being 49.4% ($=e^{0.402} - 1$) higher in 2010–2012 than in 2006–2008 and by as much as 155.5% ($=e^{0.938} - 1$) in 2018. Compared with regular check-up, prenatal/family planning visits and visits requiring medical treatment increased the odds of incurring positive OOP expenditures, whereas vaccination reduced the likelihood of incurring any OOP expenditures. By contrast, only medical treatment was a significant determinant of the level of OOP expenditures. The likelihood of having positive OOP expenditures and its magnitude varied positively with age. Being married reduced the odds of incurring any OOP expenditures, but it increased the level of OOP expenditures. Gender had no significant influence on any of the two outcome variables. Both the likelihood of having any OOP expenditures and its amount were positively associated with the level of educational attainment, although only having primary or secondary education was statistically significant. The household consumption expenditure gradients were more pronounced for the amount of OOP expenditures incurred than for the likelihood of having positive OOP expenditures. Compared with patients in the poorest expenditure quintile, patients in the bottom second and third expenditure quintiles spent 24.1% ($=e^{0.216} - 1$) and 41.7% ($=e^{0.349} - 1$) more, respectively, and those in the top fourth and fifth expenditure quintiles spent as much as 71.2% ($=e^{0.538} - 1$) and 116.1% ($=e^{0.77} - 1$) more. Similarly, the regional gradients in the amount of OOP expenditures were more pronounced for those incurring any OOP expenditures than those having zero OOP

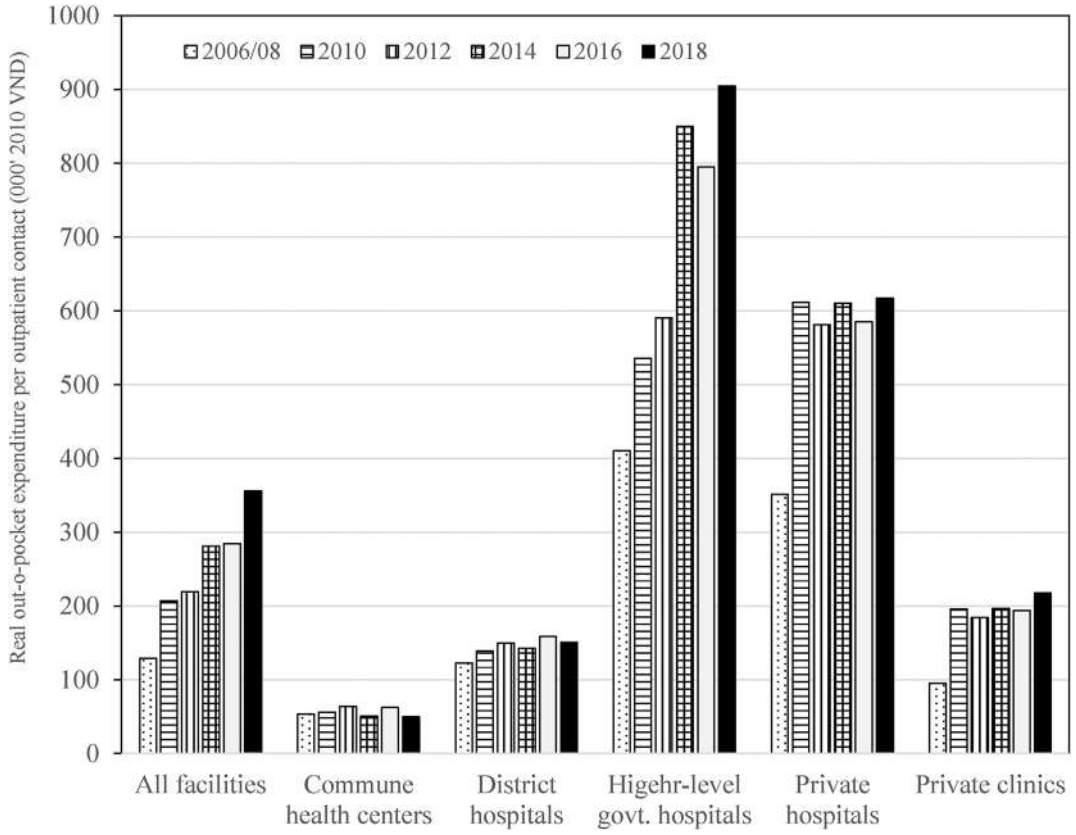


Fig. 1. Real out-of-pocket expenditure per outpatient contact by the type of facility and year (000' 2010 VND).

expenditure. Ethnicity was positively associated with the likelihood of incurring any OOP expenditures and its amount.

Controlling for the types of health facility changed the direction of the time trend gradients in the likelihood of having positive OOP expenditures, with the odds of having positive OOP expenditures being 14.9% lower in 2010–2012 than in 2006–2008 and by 20.3% in 2018, although these differences were not statistically significant. However, the results for year dummies in the GLM model remained positive and highly significant but smaller in magnitude by about 63%–65%. Seeking care at higher level government hospital rather than at district hospitals increased the likelihood of having positive OOP expenditures by 1.8 times and by as much as 6.8 times for visits at private clinics while seeking care at CHCs reduced the likelihood of having positive OOP expenditures by 70.5%. Compared with district hospitals, seeking care at private clinics increased the amount of OOP expenditures by 28.1% ($=e^{0.248}-1$) and by as much as 212.9% ($=e^{1.137}-1$) and 227.4% ($=e^{1.186}-1$) for contacts at private hospitals and higher level government hospitals, respectively. By contrast, seeking care at CHCs reduced the amount of OOP expenditures by 52.2% ($=e^{-0.739}-1$). The socioeconomic gradients in both parts of the model were attenuated when accounting for the types of health facilities contacted.

The results of the interaction terms between the types of health facility and year dummies are reported for the second part of the model in Table 3 under Model 3. Consistent with the descriptive findings, the results indicate wide variations in the relative cost of accessing an outpatient care among health facilities over the period under study, with the cost inflation being more pronounced for care sought at higher level government hospitals and private hospitals than at district hospitals (reference category). For example, an outpatient contact at higher level government hospitals that cost the insured 76.2% ($=e^{0.567}-1$) more than at district hospitals in 2006–2008 cost 134.4% ($=e^{0.852}-1$) more in 2010–2012 and as much as 362.7% ($=e^{1.532}-1$) more in 2018. By contrast, the estimated interaction terms for CHCs indicated little variations in the relative cost of accessing these facilities. Given the small numbers of outpatient contacts at private facilities in 2006–2008 and private hospitals in 2010–2012, the results for these facilities should be interpreted with caution.

The results of the interaction terms between each household consumption expenditure quintile and each type of health facility are reported for the second part of the model in Table 3 under Model 4. The expenditure quintile gradients in OOP spending were more pronounced for contacts involving higher level government hospitals and private health facilities than district hospitals and CHCs. For example, patients in the top expenditure quintile spent 65.2% ($=e^{0.502}-1$) more than those in the bottom expenditure quintile for a contact at higher level government hospitals and by as much as 88.6% ($=e^{0.613}-1$) more for a contact at private clinics. By contrast, patients in the top expenditure quintile spent 29.9% ($=e^{0.261}-1$) more for a contact at district hospitals.

Discussion

Using an appropriate TMP model and seven biennial waves from the VHLSS, this article assessed temporal trends in OOP expenditures per outpatient contact and its variation across various health facilities among the insured rural residents between 2006 and 2018. The results showed an upward trend in the likelihood of having positive OOP expenditures and its amount. The amount of OOP expenditures incurred by the insured per outpatient contact was 40.3% higher in 2010–2012 than in 2006–2008 and by as much as 155.5% higher in 2018. These high-cost pressures were attenuated by 63%–65% when accounting for the types of health facility

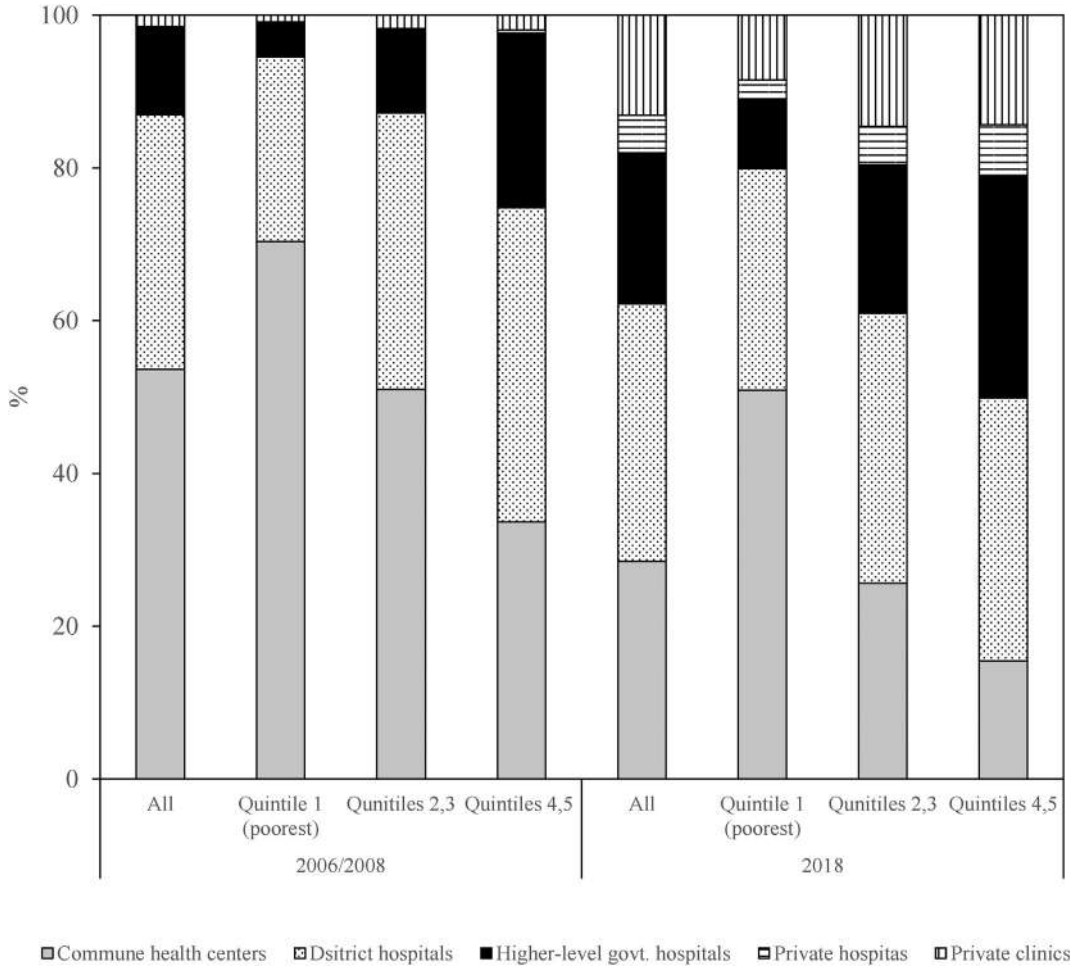


Fig. 2. Outpatient contacts by the type of health facility and household expenditure quintile, 2006–2018.

contacted. The increases in relative cost of an outpatient contact were more pronounced for contacts involving higher level government and private hospitals than other providers. The economic gradients were also more pronounced for visits at higher level government hospitals and private clinics than at lower level public health facilities.

The overall steep trend gradient in the amount of OOP expenditures and its variations across health facilities reflects a combination of factors, including changes in the outpatient care utilization pattern, the government policy of hospital autonomy, fiscal consolidation of the early and mid-2010s, and the subsequent shift toward full cost recovery for curative services provided by public health facilities and large increases in administrative prices used by VSS to reimburse providers.[17,38,39,48,49] As our descriptive statistics indicate, changes in the pattern of outpatient care utilization over the study period was large, as the utilization pattern shifted steadily toward the higher level government hospitals and private health facilities, away from CHCs. These changes in the pattern of outpatient care utilization are attributed to Vietnam's hospital-centered and fragmented delivery system, where lack of trust in the lower level public health facilities to provide proper diagnostic and quality of care often leads patients to bypass these facilities and seek care at higher level government hospitals despite substantially higher co-insurance and inconvenience.[39,49–51] The rate of self-referrals is reported to be about 41.9% at provincial hospitals, 59% at central hospital, and as high as 93.5% at the specialized hospitals/institutions (pediatrics and obstetrics).[51] For patients who are referred upward to higher level government hospitals, they are often retained at the hospitals.[39] According to one estimate, about 54%–65% of those cases presented at the higher level government hospitals could have been treated at lower level.[52] In addition to higher cost and inconveniences, patients bypassing CHCs and seeking care at higher level government hospitals and private facilities may also miss out on the promotive and preventive services that they can obtain at CHCs.[50] The large volume of patients seeking outpatient care at higher level hospitals and the resulting long waiting times and short consultation times can also have an adverse effect on doctor–patient relationship[29] and the quality of care.[53]

The government policy of hospital autonomy has steadily expanded the operational and financial autonomy of public hospitals, particularly higher level public hospitals, allowing them to mobilize financial resources from private investors, including medical staff, for the upgrade of medical equipment, and provide "elective" services that are more timely, convenient and perceived to be of higher quality.[17,55,56] The risk of over provision of high-tech health services is particularly large in an environment where providers are mainly reimbursed through an open-ended fee-for-service system, and where there is no effective oversight of providers by purchasers (VSS).[17,38,39,54,55,57] The available evidence from other settings also suggests that greater hospital autonomy may lead to an increased number of diagnostic services, provision of more profitable and specialized services.[18,58,59] The finding that the cost of accessing outpatient care grew faster for contacts at higher level

Table 2
Econometric results for the real out-of-pocket expenditures per outpatient contact.

Independent variables	Model 1				Model 2			
	Part 1 (Logit)		Part 2 (GLM)		Part 1 (Logit)		Part 2 (GLM)	
	Coef.	(95% CI)	Coef.	(95% CI)	Coef.	(95% CI)	Coef.	(95% CI)
Year dummies[a]								
Dum_2006-2008 (ref.)								
Dum_2010-2012	1.129	(0.977, 1.304)	0.402***	(0.342, 0.462)	0.851**	(0.732, 0.991)	0.151***	(0.084, 0.217)
Dum_2014-2016	1.192**	(1.032, 1.376)	0.644***	(0.582, 0.707)	0.838**	(0.719, 0.977)	0.225***	(0.156, 0.295)
Dum_2018	1.316***	(1.122, 1.543)	0.938***	(0.866, 1.010)	0.797**	(0.669, 0.948)	0.339***	(0.261, 0.416)
Types of health facility								
Commune health center					0.295***	(0.266, 0.327)	−0.739***	(−0.797, −0.681)
District hospital (ref.)								
Higher-level govt. hospital					2.847***	(2.419, 3.350)	1.186***	(1.118, 1.255)
Private hospital					5.153***	(3.307, 8.029)	1.137***	(1.036, 1.239)
Private clinic					7.841***	(5.967, 10.305)	0.248***	(0.181, 0.315)
Types of outpatient visit								
Check-ups (ref.)								
Vaccination	0.294***	(0.253, 0.343)	−0.441***	(−0.614, −0.267)	0.558***	(0.475, 0.655)	0.002	(−0.140, 0.145)
Medical treatment	1.797***	(1.633, 1.978)	0.173***	(0.117, 0.230)	1.646***	(1.488, 1.821)	0.125***	(0.070, 0.180)
Family planning	1.935***	(1.573, 2.381)	−0.102	(−0.250, 0.047)	1.728***	(1.380, 2.165)	0.011	(−0.128, 0.149)
Age	1.029***	(1.022, 1.035)	0.015***	(0.011, 0.018)	1.016***	(1.016, 1.029)	0.007***	(0.004, 0.010)
Age-squared	1.000***	(0.100, 1.00)	0.000***	(0.000, 0.000)	1.000***	(1.000, 1.000)	0.000*	(0.000, 0.000)
Gender	0.960	(0.910, 1.013)	−0.018	(−0.058, 0.022)	0.971	(0.918, 1.027)	0.005	(−0.031, 0.041)
Marital status	0.874***	(0.758, 1.009)	0.097***	(0.039, 0.154)	0.801***	(0.692, 0.928)	0.056**	(0.002, 0.109)
Level of educational attainment[b]								
Below primary (ref.)								
Primary or secondary	1.251***	(1.145, 1.367)	0.064**	(0.009, 0.118)	1.172***	(1.070, 1.285)	0.018	(−0.032, 0.068)
High school and higher	1.095	(0.953, 1.259)	−0.014	(−0.088, 0.059)	0.923	(0.798, 1.068)	−0.111***	(−0.178, −0.045)
Household expenditure quintiles								
Quintile 1 (the poorest) (ref.)								
Quintile 2	1.389***	(1.247, 1.548)	0.216***	(0.147, 0.284)	1.223***	(1.903, 1.368)	0.097***	(0.032, 0.162)
Quintile 3	1.611***	(1.431, 1.813)	0.349***	(0.274, 0.423)	1.313***	(1.160, 1.487)	0.169***	(0.103, 0.236)
Quintile 4	1.826***	(1.591, 2.095)	0.538***	(0.462, 0.613)	1.369***	(1.181, 1.586)	0.282***	(0.211, 0.353)
Quintile 5	1.956***	(1.652, 2.316)	0.770***	(0.685, 0.856)	1.297***	(1.084, 1.553)	0.435***	(0.356, 0.513)
Ethnicity	1.478***	(1.259, 1.736)	0.131***	(0.041, 0.220)	1.116	(0.947, 1.316)	−0.060	(−0.140, 0.019)
Regions								
Red River Delta (ref.)								
Northeast	0.717***	(0.606, 0.848)	−0.042	(−0.136, 0.052)	0.863	(0.721, 1.032)	0.002	(−0.090, 0.095)
Northwest	0.758*	(0.573, 1.002)	−0.020	(−0.175, 0.135)	0.925	(0.688, 1.245)	0.049	(−0.096, 0.194)
North Central Coast	0.931	(0.791, 1.094)	−0.107**	(−0.196, −0.018)	1.063	(0.890, 1.270)	−0.030	(−0.118, 0.059)
South Central Coast	0.661***	(0.549, 0.795)	−0.377***	(−0.467, −0.286)	0.638***	(0.525, 0.776)	−0.506***	(−0.595, −0.417)
Central Highlands	0.974	(0.789, 1.204)	−0.261***	(−0.371, −0.151)	1.083	(0.863, 1.359)	−0.334***	(−0.440, −0.227)
Southeast	1.118	(0.886, 1.411)	−0.396***	(−0.49, −0.299)	1.087	(0.847 1.397)	−0.479***	(−0.571, −0.386)
Mekong Delta	1.345***	(1.145, 1.580)	−0.687***	(−0.766, −0.608)	1.561***	(1.310, 1.861)	−0.792***	(−0.870, −0.715)
Outliers			2.731***	(2.647, 2.815)			2.303***	(2.200, 2.405)
n	39,762		29,318					

CI, confidence interval.
*$P < 0.1$, **$P < 0.05$, ***$P < 0.01$.
[a] Testing for equality among the year dummies suggested that certain year dummies could be aggregated.
[b] Testing for equality among the level of educational attainment dummies suggested that certain educational dummies could be aggregated.

government hospitals than district hospitals may reflect an increase in the intensity of care provided by these facilities over time, at least to the extent to which the recent increases in administrative prices increase the incentives to provide more lucrative procedures as well as "on-demand" services for which patients who have the necessary resources pay additional fees.[17,38,39,55] Our results on the influence of socio-economic variables are generally in line with those reported elsewhere. The finding that the likelihood of incurring positive OOP expenditures or the amount of OOP expenditure both vary positively with the household economic status is consistent with findings from other LMICs.[12,43,60,61] The finding that household expenditure gradients in the likelihood of incurring positive OOP expenditures or the amount of OOP expenditures are weakened when accounting for the types of health facility contacted indicates that household economic status may affect OOP spending directly as well as indirectly through provider choice.[61] Finally, the finding that ethnicity has little influence on likelihood of incurring positive OOP expenditures or the amount of OOP expenditure once we control for the types of health facility may reflect Vietnam's regional inequities in the availability and quality of both public and private providers.[39,50,62,63] The mountainous and remote regions with a higher representation of ethnic minorities tend to have a coverage rate and quality of care far below the other regions.[50,63] In our sample, contacts at CHCs in Vietnam's mountainous regions in 2018 accounted for over 57% of all outpatient contacts by the ethnic minorities vs 22.4% for ethnic majority.

These findings have important policy implications for healthcare financing and delivery over the transition path to UHC in Vietnam and other developing nations. First, the finding that the steep upward trend in OOP expenditure was largely driven by the shift in the pattern of healthcare utilization, from the lower level public health facilities toward higher level government hospitals and private health facilities, suggests that efforts should be directed at reforming Vietnam's hospital-centric and fragmented delivery system as a way of containing costs and ensuring the financial sustainability of SHI system. Primary health care is indeed widely recognized as the most cost-effective and equitable route to making progress toward

Table 3
Econometric results for the interaction terms.[a]

Independent variables	Part 2 (GLM)			
	Model 3		Model 4	
	Coef.	(95% CI)	Coef.	(95% CI)
Year dummy × type of health facility[b]				
dum_2006-2008 × commune health center	−0.646***	(−0.711, −0.580)		
dum_2006-2008 × district hospital (ref.)				
dum_2006-2008 × higher-level govt. hospital	0.567***	(0.465, 0.688)		
dum_2006-2008 × private hospital	0.662***	(0.216, 1.108)		
dum_2006-2008 × private clinic	−0.194*	(−0.404, 0.017)		
dum_2010-2012 × commune health center	−0.726***	(−0.822, −0.630)		
dum_2010-2012 × district hospital (ref.)				
dum_2010-2012 × higher-level govt. hospital	0.852***	(0.0769, 0.935)		
dum_2010-2012 × private hospital	0.907***	(0.752, 1.063)		
dum_2010-2012 × private clinic	0.166***	(0.059, 0.273)		
dum_2014-2016 × commune health center	−0.889***	(−0.988, −0.790)		
dum_2014-2016 × district hospital (ref.)				
dum_2014-2016 × higher-level govt. hospital	1.264***	(1.179, 1.349)		
dum_2014-2016 × private hospital	1.077***	(0.945, 1.209)		
dum_2014-2016 × private clinic	0.176***	(0.072, 0.281)		
dum_2018 × commune health center	−1.038***	(−1.199, −0.878)		
dum_2018 × district hospital (ref.)				
dum_2018 × higher-level govt. hospital	1.532***	(1.420, 1.645)		
dum_2018 × private hospital	1.205***	(1.025, 1.384)		
dum_2018 × private clinic	0.442***	(0.304, 0.579)		
Household consumption expenditure quintile × type of health facility[c]				
quintile 1 × commune health center (ref.)				
quintile 2 × commune health center			0.125**	(0.020, 0.231)
quintile 3 × commune health center			0.200***	(0.095, 0.305)
quintile 4 × commune health center			0.278***	(0.137, 0.419)
quintile 5 × commune health center			0.592***	(0.384, 0.799)
quintile 1 × district hospital (ref.)				
quintile 2 × district hospital			0.053	(−0.051, 0.157)
quintile 3 × district hospital			0.094*	(−0.012, 0.199)
quintile 4 × district hospital			0.206***	(0.097, 0.315)
quintile 5 × district hospital[c]			0.261***	(0.138, 0.384)
quintile 1 × higher-level govt. hospital (ref.)				
quintile 2 × higher-level govt. hospital			0.159**	(0.022, 0.296)
quintile 3 × higher-level govt. hospital			0.215***	(0.077, 0.353)
quintile 4 × higher-level govt. hospital			0.344***	(0.212, 0.476)
quintile 5 × higher-level govt. hospital			0.502***	(0.367, 0.636)
quintile 1 × private clinic (ref.)				
quintile 2 × private clinic			0.122*	(−0.021, 0.265)
quintile 3 × private clinic			0.289***	(0.140, 0.438)
quintile 4 × private clinic			0.362***	(0.210, 0.514)
quintile 5 × private clinic			0.613***	(0.418, 0.807)

CI, confidence interval.
*$P < 0.1$, **$P < 0.05$, ***$P < 0.01$.
[a] Other regressors for each model correspond to those listed in Table 2.
[b] Testing for equality among the interaction terms suggested that certain year dummies could be aggregated.
[c] Given the small number of poor patients who used private hospitals, these facilities were not interacted with household consumption expenditure quintiles.

UHC.[20,22,24,64] Sufficient investments to ensure that infrastructure and equipment availability and readiness to deliver hypertension- and diabetes-related services, including community-based screening, management, and monitoring, are predicted to increase the utilization rate of CHCs by as much as 3.3–3.7 times above their current level.[50] Second, the finding that the cost inflation was more pronounced for care sought at higher level government hospitals than at other public health facilities calls for policy measures that would make higher level public hospitals more accountable, including a stronger regulatory system and reform of hospital governance, a shift from fee-for-service to other forms of provider payment, more effective oversight of providers by purchasers, and an overall plan for hospital capacity and high-cost equipment.[17,19,38,65,66]

Although the quality of the data in this study is quite high for a developing country, common data limitations remain. The data on OOP expenditures are subject to recall errors. The estimated coefficients may be subject to endogeneity bias if the choice of provider is endogenous to the process of utilization of outpatient care. Any recall error and endogeneity bias would likely be similar across time and should not affect the trend analysis of the likelihood of having positive OOP expenditures and its amount. Moreover, in the absence of data on the types of illness and its severity, the reported results may be subject to omission bias. However, the inclusion of four broad types of outpatient visits and the types of health facility among the covariates may act as a proxy for patient case-mix, at least to the extent that higher level health facilities tend to treat more severe and complex conditions.

Author statements

Ethical approval

This study used de-identified data from Vietnam's General Statistics Office; therefore, no ethical approval was required.

Funding

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

Competing interests

The authors declare that they have no conflict of interest.

Availability of data and material

Seven biennial waves of Vietnam's Household Living Standard Survey for the period 2006–2018 were used for this study. Vietnam's Household Living Standard Survey is available from General Statistical Office of Vietnam, https://www.gso.gov.vn/en/homepage.

References

1. WHO (World Health Organization). *The world health report: health systems financing: the path to universal coverage*. Geneva: World Health Organization; 2010.
2. Atun R, de Andrade LOM, Almeida G, Cotlear D, Dmytraczenko T, Frenz P, et al. Health-system reform and universal health coverage in Latin America. *Lancet* 2015;**385**:1230–47. https://doi.org/10.1016/S0140-6736(14)61646-9.
3. Erlangga D, Suhrcke M, Ali S, Bloor K. The impact of public health insurance on health care utilisation, financial protection and health status in low- and middle-income countries: a systematic review. *PLoS One* 2019;**14**:e0219731. https://doi.org/10.1371/journal.pone.0219731.
4. Giedion U, Alfonso EA, Diaz Y. *The impact of universal coverage schemes in the developing world: a review of the existing evidence. Universal Health Coverage (UNICO) Studies Series*. Washington DC: World Bank; 2013.
5. Gilardino RE, Valanzasca P, Rifkin SB. Has Latin America achieved universal health coverage yet? Lessons from four countries. *Arch Public Health* 2022;**80**:38. https://doi.org/10.1186/s13690-022-00793-7.
6. Lagomarsino G, Garabrant A, Adyas A, Muga R, Otoo N. Moving towards universal health coverage: health insurance reforms in nine developing countries in Africa and Asia. *Lancet* 2012;**380**:933–43. https://doi.org/10.1016/S0140-6736(12)61147-7.
7. Marten R, McIntyre D, Travassos C, Shishkin S, Longde W, Reddy S, et al. An assessment of progress towards universal health coverage in Brazil, Russia, India, China, and South Africa (BRICS). *Lancet* 2014;**384**:2164–71. https://doi.org/10.1016/S0140-6736(14)60075-1.
8. Reich MR, Harris J, Ikegami N, Maeda A, Cashin C, Araujo EC, et al. Moving towards universal health coverage: lessons from 11 country studies. *Lancet* 2016;**387**:811–6. https://doi.org/10.1016/S0140-6736(15)60002-2.
9. Chankova S, Sulzbach S, Diop F. Impact of mutual health organizations: evidence from West Africa. *Health Policy Plan* 2008;**23**:264–76. https://doi.org/10.1093/heapol/czn011.
10. Vu PH, Sepehri A, Tran LTT. Trends in out-of-pocket expenditure on facility-based delivery and financial protection of health insurance: findings from Vietnam's Household Living Standard Survey 2006-2018. *Int J Health Econ Manag* 2022 Apr 13. https://doi.org/10.1007/s10754-022-09330-z.
11. Wagstaff A, Yu S. Do health sector reforms have their intended impacts? The World Bank's Health VIII Project in Gansu Province, China. *Health Econ* 2007;**26**:505–35. https://doi.org/10.1016/j.jhealeco.2006.10.006.
12. Wang W, Temsah G, Carter E. *Levels and determinants of out-of-pocket health expenditures in the Democratic Republic of the Congo, Liberia, Namibia, and Rwanda. DHS analytical studies 59*. Rockville, Maryland: ICF International; 2016.
13. Wang H, Yip W, Zhang L, Hsiao W. Community-based health insurance in poor rural China: the distribution of net benefits. *Health Policy Plan* 2005;**20**:366–74. https://doi.org/10.1093/heapol/czi045.
14. Yang W, Wu X. Paying for outpatient care in rural China: cost escalation under China's New Co-operative Medical Scheme. *Health Policy Plan* 2015;**30**:187–96. https://doi.org/10.1093/heapol/czt111.
15. Eggleston K, Li L, Qingyue M, Lindelow M, Wagstaff A. Health service delivery in China: a literature review. *Health Econ* 2008;**17**:149–65. https://doi.org/10.1002/hec.1306.
16. Li C, Chen Z, Khan MM. Bypassing primary care facilities: health-seeking behavior of middle age and older adults in China. *BMC Health Serv Res* 2021;**21**:895. https://doi.org/10.1186/s12913-021-06908-0.
17. Somanathan A, Tandon A, Dao HL, Huong L, Hurt KL, Kari L, et al. *Moving toward Universal Coverage of Social Health Insurance in Vietnam: Assessment and Options. Directions in Development – Human Development*. Washington, DC: World Bank; 2014.
18. Wagstaff A, Lindelow M. Can insurance increase financial risk? The curious case of health insurance in China. *J Health Econ* 2008;**27**:990–1005. https://doi.org/10.1016/j.jhealeco.2008.02.002.
19. Yip W, Hsiao W. Harnessing the privatization of China's fragmented health-care delivery. *Lancet* 2014;**384**:805–18. https://doi.org/10.1016/s0140-6736(14)61120-x.
20. Bitton A, Ratcliffe HL, Veillard JH, Kress DH, Barkley S, Kimball M, et al. Primary health care as a foundation for strengthening health systems in low- and middle-income countries. *J Gen Intern Med* 2017;**32**:566–71. https://doi.org/10.1007/s11606-016-3898-5.
21. Hanson K, Brikci N, Erlangga D, Alebachew A, De Allegri M, Balabanova D, et al. The Lancet Global Health Commission on financing primary health care: putting people at the centre. *Lancet Glob Health* 2022;**10**:e715–72. https://doi.org/10.1016/S2214-109X(22)00005-5.
22. Pettigrew LM, Maeseneer JD, Anderson M-IP, Essuman A, Kidd MR, Haines A. Primary health care and the sustainable development goals. *Lancet* 2015;**386**:2119–21. https://doi.org/10.1016/S0140-6736(15)00949-6.
23. Rao M, Pilot E. The missing link–the role of primary care in global health. *Glob Health Action* 2014;**7**:23693. https://doi.org/10.3402/gha.v7.23693.
24. Veillard J, Cowling K, Bitton A, Ratcliffe H, Kimball M, Barkley S, et al. Better measurement for performance improvement in low-and middle-income countries: the primary health care performance initiative (PHCPI) experience of conceptual framework development and indicator selection. *Milbank Q* 2017;**95**:836–83. https://doi.org/10.1111/1468-0009.12301.
25. Bell G, Macarayan EK, Ratcliffe H, Kim JH, Otupiri E, Lipsitz S, et al. Assessment of bypass of the nearest primary health care facility among women in Ghana. *JAMA Netw* 2020;**3**:e2012552. https://doi.org/10.1001/jamanetworkopen.2020.12552.
26. Liu Y, Zhong L, Yuan S, van de Klundert J. Why patients prefer high-level healthcare facilities: a qualitative study using focus groups in rural and urban China. *BMJ Glob Health* 2018a;**3**:e000854. https://doi.org/10.1136/bmjgh-2018-000854.
27. Liu Y, Kong Q, Yuan S, van de Klundert J. Factors influencing choice of health system access level in China: a systematic review. *PLoS One* 2018b;**13**:e0201887. https://doi.org/10.1371/journal.pone.0201887.
28. Rao KD, Sheffel A. Quality of clinical care and bypassing of primary health centers in India. *Soc Sci Med* 2018;**207**:80–8. https://doi.org/10.1016/j.socscimed.2018.04.040.
29. Wu D, Wang Y, Lam KF, Hesketh T. Health system reforms, violence against doctors and job satisfaction in the medical profession: a cross-sectional survey in Zhejiang Province, Eastern China. *BMJ Open* 2014;**4**:e006431. https://doi.org/10.1136/bmjopen-2014-006431.
30. Yip W, Fu H, Chen AT, Zhai T, Jian W, Xu R, et al. 10 years of health-care reform in China: progress and gaps in Universal Health Coverage. *Lancet* 2019;**394**:1192–204. https://doi.org/10.1016/S0140-6736(19)32136-1.
31. Liu M, Zhang Q, Lu M, Kwon CS, Quan H. Rural and urban disparity in health services utilization in China. *Med Care* 2007;**45**:767–74. https://doi.org/10.1097/MLR.0b013e3180618b9a.
32. Xu J, Zheng J, Xu L, Wu H. Equity of health services utilisation and expenditure among urban and rural residents under universal health coverage. *Int J Environ Res Public Health* 2021;**18**:593. https://doi.org/10.3390/ijerph18020593.
33. Karan A, Selvaraj S, Mahal A. Moving to universal coverage? Trends in the burden of out-of-pocket payments for health care across social groups in India, 1999–2000 to 2011–12. *PLoS One* 2014;**9**:e105162. https://doi.org/10.1371/journal.pone.0105162.
34. Kanmiki EW, Bawah AA, Phillips JF, Awoonor-Williams JK, Kachur SP, Asuming PO, et al. Out-of-pocket payment for primary healthcare in the era of national health insurance: evidence from northern Ghana. *PLoS One* 2019;**14**:e0221146. https://doi.org/10.1371/journal.pone.0221146.
35. Qin VM, Zhang Y, Chia KS, Zhao Y, Hulse ESG, Legido-Quigley H, et al. Temporal trends and variation in out-of-pocket expenditures and patient cost sharing: evidence from a Chinese national survey 2011-2015. *Int J Equity Health* 2021;**20**:143. https://doi.org/10.1186/s12939-021-01480-9.
36. Ta Y, Zhu Y, Fu H. Trends in access to health services, financial protection and satisfaction between 2010 and 2016: has China achieved the goals of its health system reform? *Soc Sci Med* 2020;**245**:112715. https://doi.org/10.1016/j.socscimed.2019.112715.
37. Le QN, Blizzard L, Si L, Giang LT, Neil A. The evolution of social health insurance in Vietnam and its role towards achieving universal health coverage. *Health Policy OPEN* 2020;**1**:100011. https://doi.org/10.1016/j.hpopen.2020.100011.
38. Teo HS, Bales S, Bredenkamp C, Cain J. *The Future of Health Financing in Vietnam: Ensuring Sufficiency, Efficiency, and Sustainability*. Washington, DC: World Bank; 2019.
39. World Bank. *Moving toward UHC: Vietnam - national initiatives, key challenges, and the role of collaborative activities (English). Moving toward UHC*. Washington, DC: World Bank; 2017. http://documents.worldbank.org/curated/en/464111513160157867/Moving-toward-UHC-Vietnam-national-initiatives-key-challenges-and-the-role-of-collaborative-activities. [Accessed 22 September 2021].
40. Ha NT, Anh NQ, Van Toan P, Huong NT. Health insurance reimbursement to hospitals in Vietnam: policy implementation results and challenges. *Health*

Serv Insights 2021;**14**:11786329211010126. https://doi.org/10.1177/11786329211010126.
41. Teo HS, Huong DL. *Improving efficiency in the health sector: an assessment of Vietnam's readiness for integration of care*. Washington, DC: World Bank; 2020.
42. Deb P, Norton EC. Modeling health care expenditures and use. *Annu Rev Public Health* 2018;**39**:489–505. https://doi.org/10.1146/annurev-publhealth-040617-013517.
43. Herberholz C, Phuntsho S. Medical, transportation and spiritual out-of-pocket health expenditure on outpatient and inpatient visits in Bhutan. *Soc Sci Med* 2021;**273**:13780. https://doi.org/10.1016/j.socscimed.2021.113780.
44. Mohanty SK, Srivastava A. Out-of-pocket expenditure on institutional delivery in India. *Health Policy Plan* 2013;**28**:247–62. https://doi.org/10.1093/heapol/czs057.
45. Wang Q, Fu AZ, Brenner S, Kalmus O, Banda HT, De Allegri M. Out-of-pocket expenditure on chronic non-communicable diseases in sub-Saharan Africa: the case of rural Malawi. *PLoS One* 2015;**10**:e0116897. https://doi.org/10.1371/journal.pone.0116897.
46. Deaton A. *The analysis of household surveys: a microeconometric approach to development policy*. Washington, DC: World Bank; 2019.
47. Rabe-Hesketh S, Skrondal A. *Multilevel and longitudinal modeling using Stata*. Texas: Stata Corporation; 2005.
48. Duc TN, Anh BTM, Thanh HP, Quynh AP, Huyen XC. Impact of public health insurance on out-of-pocket health expenditures of the near-poor in Vietnam. *Health Serv Insights* 2021;**14**:11786329211017411. https://doi.org/10.1177/11786329211017411.
49. Duong DB, Minh HV, Ngo LH, Ellner AL. Readiness, availability and utilization of rural Vietnamese health facilities for community based primary care of non-communicable diseases: a cross sectional survey of 3 provinces in northern Vietnam. *Int J Health Policy Manag* 2019;**8**:150–7. https://doi.org/10.15171/ijhpm.2018.104.
50. Vu LT, Bales S, Bredenkamp C. What drives utilization of primary care facilities?: evidence from a national facility survey. *Int J Healthc Manag* 2021;**14**:1107–13. https://doi.org/10.1080/20479700.2020.1752984.
51. World Bank. *Project appraisal document on health professionals education and training project*. Hanoi: World Bank; 2014. http://documents.worldbank.org/curated/en/427731468316154123/pdf/823180PAD0IDA0010Box385188B00OUO090.pdf. [Accessed 10 May 2022].
52. MOH (Ministry of Health) Vietnam, and HPG (Health Partnership Group). *Joint annual health review JAHR 2015: strengthening primary health care at the grassroots towards universal health coverage*. Hanoi: Ministry of Health; 2015.
53. Zhang L, Cheng G, Song S, Yuan B, Zhu W, He L, et al. Efficiency performance of China's health care delivery system. *Int J Health Plann Manage* 2017;**32**:254–63. https://doi.org/10.1002/hpm.2425.
54. Võ MTH, Löfgren K. An institutional analysis of the fiscal autonomy of public hospitals in Vietnam. *Asia Pac Policy Stud* 2019;**6**:90–107. https://doi.org/10.1002/app5.268.
55. Lieberman S, Wagstaff A. *Health financing and delivery in Vietnam: looking forward. World Bank/Health, Nutrition, and Population Series*. Washington, DC: World Bank; 2009.
56. MOH (Ministry of Health). *Health Strategy and Policy Institute, World Bank. Lessons for Hospital Autonomy Implementation in Vietnam from International Experience: Issues Identified from International Studies and a Public Hospital Survey in Vietnam*. Hanoi: World Bank; 2011.
57. Wagstaff A, Bales S. *The impacts of public hospital autonomization: evidence from a quasi-natural experiment. World Bank policy research working paper series 6137*. 2012. https://openknowledge.worldbank.org/handle/10986/11942. [Accessed 17 June 2014].
58. Allen P, Cao Q, Wang H. Public hospital autonomy in China in an international context. *Int J Health Plann Manage* 2014;**29**:141–59. https://doi.org/10.1002/hpm.2200.
59. Ravaghi H, Foroughi Z, Nemati A, Bélorgeot V. A holistic view on implementing hospital autonomy reforms in developing countries: a systematic review. *Health Policy Plan* 2018;**33**:1118–27. https://doi.org/10.1093/heapol/czy095.
60. Farag M, NandaKumar AK, Wallack S, Hodgkin D, Gaumer G, Erbil C. The income elasticity of health care spending in developing and developed countries. *Int J Health Care Finance Econ* 2012;**12**:45–62. https://doi.org/10.1007/s10754-012-9108-z.
61. Rousa J, Hotchkiss D. Estimation of the determinants of household health care expenditures in Nepal with controls for endogenous illness and provider choice. *Health Econ* 2003;**12**:431–51. https://doi.org/10.1002/hec.727.
62. Demombynes G, Hurt K. *Quality and equity in basic health care services in Vietnam: findings from the Vietnam District and Commune Health Facility Survey*. Washington: World Bank; 2015. https://microdata.worldbank.org/index.php/catalog/2728/download/39045. [Accessed 15 March 2021].
63. GSO (General Statistics Office) Vietnam, and UNICEF (United Nations Children's Fund). *Vietnam multiple indicator cluster survey 2014. Final Report, Hanoi, Vietnam*. 2015.
64. WHO (World Health Organization). *Primary health care on the road to universal health coverage: 2019 global monitoring report*. Geneva: World Health Organization; 2020.
65. Fang H, Eggleston K, Hanson K, Wu M. Enhancing financial protection under China's social health insurance to achieve universal health coverage. *BMJ* 2019;**365**:l2378. https://doi.org/10.1136/bmj.l2378.
66. Yip WC-M, Hsiao WC, Chen W, Hu S, Ma J, Maynard A. Early appraisal of China's huge and complex health-care reforms. *Lancet* 2012;**379**:833–42. https://doi.org/10.1016/S0140-6736(11)61880-1.

Public Health 215 (2023) 94–99

Contents lists available at ScienceDirect

Public Health

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

Original Research

Comparative analysis of risk factors for COVID-19 mortality before, during and after the vaccination programme in Mexico

G.M. Parra-Bracamonte [a, *], N. Lopez-Villalobos [b], M.A. Velazquez [c], F.E. Parra-Bracamonte [d], A.L. Perales-Torres [e], K.J. Juárez Rendón [a]

[a] Centro de Biotecnología Genómica-Instituto Politécnico Nacional, Reynosa, 87710, Mexico
[b] School of Agriculture and Environment, Massey University, Palmerston North, 4442, New Zealand
[c] School of Natural & Environmental Sciences, Newcastle University, Newcastle Upon Tyne, NE1 7RU, United Kingdom
[d] Cardiology Clinical Care, Centro Médico Pensiones, Mérida, 97070, Mexico
[e] Unidad Académica Reynosa-Aztlán, Universidad Autónoma de Tamaulipas, Reynosa, 88710, Mexico

ARTICLE INFO

Article history:
Received 24 June 2022
Received in revised form
18 November 2022
Accepted 15 December 2022
Available online 21 December 2022

Keywords:
Comorbidities
SARS-CoV-2
Odds ratio
Vaccination

ABSTRACT

Objectives: The purpose of this study was to compare case fatality rates (CFRs) and odds for mortality by risk factors of patients with COVID-19 in Mexico, before, during and after the implementation of the national COVID-19 vaccination programme.
Study design: A large database including COVID-19 monitoring cases was used to perform an observational retrospective study.
Methods: The Chi-squared test and multivariate logistic regression analyses were applied to data from COVID-19-positive patients in Mexico. Data were analysed over 3 years, 2020, 2021 and 2022, corresponding with pre-, during and post-vaccination periods. The unadjusted odds ratios and 95% confidence interval were used to estimate the risk factors for COVID-19 mortality in each of the years.
Results: Statistically significant differences in CFR and odds ratio were found in the studied years, favouring postvaccination period. Significant changes in CFR by age, sex and main comorbidities indicated changes in the epidemic dynamics after the implementation of the COVID-19 vaccination campaign. The likelihood of death increased for hospitalised cases and for patients who were middle-aged or older in 2021 and 2022, whereas the odds of death associated with sex and comorbidities remained similar or reduced over the 3 years.
Conclusions: Implementation of the COVID-19 vaccination programme during 2021 showed positive consequences on CFR. The increased odds of dying in hospitalised patients are likely to be due to the unvaccinated proportion of patients.

Introduction

In December 2019, a COVID-19 of the respiratory tract characterised by a severe acute respiratory syndrome, caused by a betacoronavirus named SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2),[1,2] emerged in Wuhan City, Hubei province in China.[3] Despite rigorous global containment and quarantine efforts, the COVID-19 outbreak become a global pandemic. Currently, with a global vaccination strategy in all continents, the disease has resulted in almost 500 million positive cases and around 6.2 million deaths.[4] In Mexico, the first cases of COVID-19 were registered in January 2020, and official confirmation was published in late February 2020. On 20 April 2020, the Mexican government officially declared that the country had entered phase 3 of the COVID-19 pandemic (i.e. epidemic phase). Following three major peaks of infections, the epidemic has now decreased to an incidence rate of 3.3 cases per 100,000 inhabitants.[5]

Mexico has population of almost 130 million, of whom 52% and 48% are female and male, respectively,[6] with a high prevalence of comorbidities, such as hypertension, obesity and diabetes, which are considered the risk factors for death.[7] After a year of implementing the national COVID-19 vaccination programme that formerly prioritised front-line health workers and elderly patients,

* Corresponding author. Centro de Biotecnología Genómica-IPN, Blvd. del Maestro SN, Esq. Elías Piña, Col. Narciso Mendoza, Reynosa, C.P. 88710, Tamaulipas, Mexico. Tel.: +52 8999243627x87709.
E-mail address: gparra@ipn.mx (G.M. Parra-Bracamonte).

https://doi.org/10.1016/j.puhe.2022.12.006
0033-3506/

with or without main comorbidities (i.e. hypertension, diabetes and obesity), it is estimated that >65% of the national population is now fully vaccinated.[8,9]

Although COVID-19 outcomes and risk factors associated with death have now been documented for Mexico,[10–12] there is a lack of national epidemiological analyses after the implementation of the COVID-19 vaccination campaign, including the emergence of new viral variants resulting in important peaks of the disease.[13,14] Hence, the objective of the present study was to compare case fatality rates (CFRs) and the risks of mortality for COVID-19-positive patients during three 1-year periods corresponding with the time before, during and after (i.e. 2020, 2021 and 2022, respectively) the implementation of the national COVID-19 vaccination programme in Mexico.

Methods

Data source

A database including COVID-19 monitoring cases was downloaded from the open data source of the Epidemiologic Surveillance Source of Respiratory Viral Diseases (Sistema de Vigilancia Epidemiológica de Enfermedades Respiratorias Virales) that contained information from 475 monitoring units across the country from the public and private health sectors. Positive cases were extracted and edited. Data from 6,657,667 patients diagnosed as positive for COVID-19, from the first positive case registered on 13 January to 24 october 2022 (database accessed on 25 october 2022) were analysed. The data were divided into three 1-year data sets, as follows: C2020, cases until 23 December 2020, a day before the start date of the national vaccination programme;[15] C2021, cases from 24 December 2020 to 23 December 2021; and C2022, cases from 24 December 2021 to 24 October 2022. All COVID-19-positive cases were diagnosed using real-time polymerase chain reaction and were officially registered by the National Network for Epidemiological Surveillance (Red Nacional de Laboratorios de Vigilancia Epidemiologica), recognised by the Institute of Epidemiological Diagnosis and Reference (Instituto de Diagnóstico y Referencia Epidemiológicos, InDRE).

Each patient record included information on age, sex, smoking habits, exposure history, comorbidity traits and clinical care management. Sex was recorded as male or female. The following characteristics were recorded as 'yes' or 'no': smoking habits, hospitalisation, endotracheal intubation, intensive care unit (ICU) admission, hypertension, obesity, cardiopathy, pneumonia, chronic obstructive pulmonary disease (COPD), asthma, immunosuppression, chronic kidney disease (CKD) and other complications.[10] CFR for each year was calculated by dividing the number of deaths from COVID-19 by the number of individuals diagnosed with COVID-19, and the resulting ratio was then multiplied by 100 to be expressed as a percentage. CFRs were calculated for patient characteristics.

Statistical analyses

All statistical analyses were performed using SAS v.9.4 (SAS Institute Inc., Cary, NC, USA). Differences in CFR by year, clinical characteristics and comorbidities were examined by a Chi-squared test using the FREQ procedure, and the differences were confirmed by an exact logistic regression assuming a logit binomial distribution of data using the GENMOD procedure. The unadjusted odds ratio and 95% confidence interval for the different levels of risk factors of COVID-19 were modelled with a multivariate logistic regression model that included the effects of age, sex, smoking habits, patient hospitalisation and comorbidity traits. Comorbidities included hypertension, obesity, pneumonia, COPD, asthma, immunosuppression, CKD and other complications, using the LOGISTIC procedure. For the current analysis, age was classified into 18 groups of 5-year intervals, from 0 to up to >84 years old. Statistical significance was set at <0.01.

Results

The frequencies and CFR for each year group are presented in Table 1. CFRs were significantly different among the three year groups ($P < 0.0001$). The C2022 group showed the highest frequency of positive cases; however, the CFR was lower than in the C2020 group. In terms of age groups, the CFR in C2020 was only noticeably different to C2021 for some elderly age groups (>60 years); however, a progressive decrease in CFR was observed over the study period, largely during C2022 among all age groups. CFR across sex, smoking habits and comorbidities from C2020 to C2022 significantly decreased ($P < 0.0001$). Interestingly, the absolute number of hospitalisations and related traits remained fairly similar during C2020 and C2021, with a slight reduction in C2022, yet their frequencies were statistically different ($P < 0.0001$).

Multivariate odds ratios of mortality were estimated for age, sex, smoking habits, hospitalisation admission characteristics and comorbidities (Table 2). The results revealed an age-dependent increase in the odds of mortality for individuals middle-aged and older during C2021 and C2022, where the likelihood of death was twice as high compared with C2020 (Table 2). With a significant and slight reduction over the 3 years, male patients had higher risk of death than female patients.

A significant risk was associated with smoking habits, where smokers showed slightly lower odds of mortality than non-smokers. In terms of comorbidities, a slight reduction from C2020 to partial 2022 year was observed; however, the presence of comorbidities significantly increased the odds of mortality in all years analysed, with the exception of COPD, which did not influence the odds during C2021, and obesity in C2022 (Table 2).

The most significant change observed in the years analysed was the substantial increased likelihood of death associated with hospitalisation – from 17.4 to 82.7 times higher among inpatients compared with non-hospitalised COVID-19-positive patients in C2020 and C2021, respectively. The odds of mortality further increased to 146.1 in C2022 in COVID-19 hospitalised patients, an eight-fold increase compared with C2020.

Discussion

In this study, three 1-year groups were analysed from a large data set of registered COVID-19-positive cases (N = 6,657,667) in Mexico from 13 January 2020 to 24 October 2022. The odds of death related to age, sex and comorbidities in COVID-19 patients from Mexico have been previously discussed.[10–12] In general, most of the estimated odds for mortality remained similar among the year groups, and previous risk factors associated with COVID-19 were confirmed in the present analysis.[10] For instance, comorbidities previously reported as important by their odds of mortality, such as hypertension, diabetes, obesity and CKD,[10] maintained similar and significant odds in the year groups assessed in the present study, highlighting the relevance of continuous monitoring of vulnerable populations during vaccination campaigns. However, the comparative analysis of year groups revealed the following three main findings that are relevant to discuss: (1) a reduction of total CFR over the 3-year study period; (2) increasing odds of mortality in older ages during the second- and third-year group; and (3) the dramatic and sustained increase in odds of mortality in hospitalised patients.

Table 1
Frequencies and case fatality rates (CFR) in COVID-19-positive during 2020, 2021 and 2022 in Mexico according to risk factors.

Risk factor	2020				2021				2022			
	Non-survivors	Survivors	Total	CFR (%)	Non-survivors	Survivors	Total	CFR (%)	Non-survivors	Survivors	Total	CFR (%)
n	138,789	1,258,207	1,396,996	9.9	140,973	2,171,003	2,311,976	6.1	23,841	2,924,854	2,948,695	0.8
Age (years)												
0–4	262	7127	7389	3.5	219	20,479	20,698	1.1	181	36,613	36,794	0.5
5–9	48	7571	7619	0.6	55	30,085	30,140	0.2	41	49,835	49,876	0.1
10–14	67	14,411	14,478	0.5	101	58,700	58,801	0.2	52	78,243	78,295	0.1
15–19	166	34,475	34,641	0.5	271	113,351	113,622	0.2	91	101,414	101,505	0.1
20–24	501	90,270	90,771	0.6	732	231,006	231,738	0.3	129	288,240	288,369	0.0
25–29	1096	146,100	147,196	0.7	1832	294,145	295,977	0.6	231	379,891	380,122	0.1
30–34	2101	153,417	155,518	1.4	2908	264,796	267,704	1.1	319	376,921	377,240	0.1
35–39	3347	147,909	151,256	2.2	5031	244,947	249,978	2.0	422	334,645	335,067	0.1
40–44	5754	137,690	143,444	4.0	6484	207,504	213,988	3.0	513	305,582	306,095	0.2
45–49	9539	138,085	147,624	6.5	9977	200,450	210,427	4.7	871	285,337	286,208	0.3
50–54	12,772	116,093	128,865	9.9	11,869	153,231	165,100	7.2	1352	238,709	240,061	0.6
55–59	16,743	92,739	109,482	15.3	15,268	117,612	132,880	11.5	1810	172,482	174,292	1.0
60–64	19,392	64,878	84,270	23.0	17,860	84,580	102,440	14.9	2227	107,617	109,844	2.0
65–69	19,544	43,796	63,340	30.9	18,584	59,079	77,663	23.0	2669	67,801	70,470	3.8
70–74	17,465	28,390	45,855	38.1	16,855	39,361	56,216	33.1	2863	43,005	45,868	6.2
75–79	13,590	17,820	31,410	43.3	14,019	25,076	39,095	43.1	3025	27,690	30,715	9.8
80–84	9332	10,020	19,352	48.2	9816	14,677	24,493	57.2	2987	16,437	19,424	15.4
>84	7070	7416	14,486	48.8	9092	11,924	21,016	46.7	4058	14,392	18,450	22.0
Sex												
Female	5,0928	639,709	690,637	7.0	56,963	1,112,215	1,169,178	5.0	9538	1,669,679	1,679,217	0.6
Male	87,861	618,468	706,329	12.0	84,010	1,058,788	1,142,798	7.0	14,303	1,255,175	1,269,478	1.0
Smoking status				0.0				0.0				0.0
Smoking	11,035	93,776	104,811	11.0	9753	126,674	136,427	7.0	1817	118,192	120,009	2.0
Non-smoking	127,754	1,164,431	1,292,185	10.0	131,220	2,044,329	2,175,549	6.0	21,938	2,795,433	2,817,371	0.8
Hospitalisation												
Non-hospitalised	11,916	1,097,953	1,109,869	1.1	5922	2,013,746	2,019,668	0.3	1903	2,865,234	2,867,137	0.1
Hospitalised	126,873	160,254	287,127	44.0	135,051	157,257	292,308	46.0	23,839	59,620	83,459	29.0
ICU	14,336	10,378	24,714	58.0	13,302	8829	22,131	60.0	1878	2493	4371	43.0
ETI	36,300	7058	43,358	84.0	25,557	4737	30294	84.0	3556	1231	4787	74.0
Pneumonia status												
With pneumonia	102,157	116,828	218,985	47.0	100,646	112,602	213,248	47.0	14,725	38,899	53,624	27.0
Without pneumonia	36,632	1,141,379	1,178,011	3.0	40,327	2,058,401	2,098,728	2.0	9115	2,873,150	2,882,265	0.3
Comorbidity				0.0				0.0				0.0
Hypertension	63,582	190,117	253,699	25.0	60,532	226,258	286,790	21.0	11,251	245,935	257,186	4.0
Obesity	32,441	190,259	222,700	15.0	28,628	208,150	236,778	12.0	2785	190,749	193,534	1.0
Diabetes	53,089	143,606	196,695	27.0	49,738	169,299	219,037	23.0	9036	164,673	173,709	5.0
Cardiopathy	7374	16,814	24,188	30.0	6329	15,778	22,107	29.0	1749	18,303	20,052	9.0
COPD	6487	10,815	17,302	37.0	5478	10,780	16,258	34.0	1505	10,518	12,023	13.0
Asthma	2640	30,922	33,562	8.0	2326	37,691	40,017	6.0	363	50,909	51,272	0.7
Immunosuppressed	3267	9813	13,080	25.0	2809	9009	11,818	24.0	946	11,637	12,583	8.0
CKD	9781	13,225	23,006	43.0	9158	13,640	22,798	40.0	2998	14,779	17,777	17.0
Other complication	7209	22,377	29,586	24.0	6557	28,341	34,898	19.0	1815	34,149	35,964	5.0

CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; ETI, endotracheal intubation; ICU, intensive care unit.
Fatality ratio (FR) were estimated by trait.

Constant reduction in CFRs in the last three years

In Mexico, the results from the first epidemiological analyses of COVID-19 were uncertain because of the reported and unprecedented increasing CFR for some vulnerable population groups.[10] The implementation of the national COVID-19 vaccination programme, initiated on 24 December 2020, was believed to be a relief to the constant increasing number of cases and mortality. The vaccination strategy was implemented in stages, with the aim to vaccinate the majority of the Mexican population by the end of 2021. The strategy prioritised, step-by-step, (1) front-line healthcare personnel, (2) the elderly population (≥60 years old), (3) those aged 50–59 years and pregnant women, (4) those aged 40–49 years and, finally (5) the rest of the population.[15] However, some amendments were necessary to cover comorbidities in vulnerable individuals and school personnel. As the programme developed, it was evident the vaccination coverage was dependent on laborious government vaccine acquisition and reduced widespread vaccine availability that delayed vaccine uptake.[14] By the end of 2021, approximately 63% of the Mexican population was vaccinated, with 56% fully vaccinated and 7.1% partially vaccinated (i.e. had only received one dose).[9] Following the fourth peak of COVID-19 infections during the early months of 2022, it was estimated that around 66% of the Mexican population had been vaccinated.[9]

Hence, considering the emergence of the SARS-CoV-2 Delta variant and the triggering of the third wave of COVID-19 in May 2021,[13,16] the current analysis strongly suggests that vaccination against COVID-19 in Mexico was an effective way to reduce CFR in 2021 and 2022. Similarly, a study in the United Kingdom, the first country that implemented a COVID-19 vaccination programme, reported that COVID-19 vaccination was effective against symptomatic disease with the Delta variant, especially after the application of two doses.[17] Moreover, Liang et al.[18] analysed data from 90 countries from November 2020 to April 2021 and estimated that a 10% increase in vaccine coverage resulted in a 7.6% reduction in COVID-19 CFR. The gradual decrease in CFR observed in the present study from 2021 to 2022 following the implementation of the national COVID-19 vaccination programme at the end of 2020 is in line with a report from China indicating that vaccination decreases the risk of developing severe COVID-19.[19] The latter study also

Table 2
Multivariate odds ratios (±95% confidence interval) for risk factors of case fatality rate in patients positive to COVID-19 in Mexico during 2020, 2021 and 2022 in Mexico.

Risk factor	2020		2021		2022	
	Multivariate odds ratio[a]	*P*-value	Multivariate odds ratio[b]	*P*-value	Multivariate odds ratio[c]	*P*-value
Age (years)						
0–4	1.0		1.0		1.0	
5–9	0.488 (0.354–0.672)	<0.0001	0.548 (0.403–0.746)	0.0001	0.602 (0.426–0.851)	0.0040
10–14	0.430 (0.325–0.570)	<0.0001	0.692 (0.541–0.884)	0.0033	0.892 (0.648–1.227)	0.4832
15–19	0.505 (0.411–0.621)	<0.0001	0.961 (0.797–1.159)	0.6780	1.419 (1.093–1.842)	0.0086
20–24	0.570 (0.486–0.670)	<0.0001	1.218 (1.039–1.428)	0.0151	1.304 (1.032–1.647)	0.0263
25–29	0.608 (0.525–0.704)	<0.0001	1.754 (1.513–2.035)	<0.0001	1.686 (1.375–2.067)	<0.0001
30–34	0.867 (0.754–0.998)	0.0472	2.345 (2.028–2.712)	<0.0001	2.208 (1.824–2.673)	<0.0001
35–39	1.127 (0.982–1.294)	0.0887	3.290 (2.850–3.797)	<0.0001	3.169 (2.637–3.808)	<0.0001
40–44	1.598 (1.395–1.831)	<0.0001	4.082 (3.539–4.708)	<0.0001	3.663 (3.062–4.383)	<0.0001
45–49	2.108 (1.842–2.413)	<0.0001	5.043 (4.376–5.811)	<0.0001	5.369 (4.529–6.365)	<0.0001
50–54	2.691 (2.352–3.078)	<0.0001	6.100 (5.294–7.029)	<0.0001	6.902 (5.851–8.141)	<0.0001
55–59	3.634 (3.178–4.157)	<0.0001	7.766 (6.741–8.946)	<0.0001	8.082 (6.869–9.509)	<0.0001
60–64	4.970 (4.345–5.684)	<0.0001	9.972 (8.657–11.488)	<0.0001	9.693 (8.248–11.392)	<0.0001
65–69	6.341 (5.543–7.254)	<0.0001	12.239 (10.623–14.102)	<0.0001	11.127 (9.479–13.062)	<0.0001
70–74	8.000 (6.990–9.156)	<0.0001	14.620 (12.684–16.851)	<0.0001	12.883 (10.975–15.123)	<0.0001
75–79	9.493 (8.287–10.875)	<0.0001	16.642 (14.429–19.194)	<0.0001	14.476 (12.334–16.991)	<0.0001
80–84	11.197 (9.757–12.850)	<0.0001	18.502 (16.018–21.370)	<0.0001	16.700 (14.220–19.611)	<0.0001
>84	12.684 (11.033–14.583)	<0.0001	21.204 (18.347–24.508)	<0.0001	19.081 (16.279–22.366)	<0.0001
Sex						
Female	1.0		1.0		1.0	
Male	1.479 (1.455–1.502)	<0.0001	1.346 (1.326–1.368)	<0.0001	1.386 (1.340–1.439)	<0.0001
Smoking status						
Non-smoking	1.0		1.0		1.0	
Smoking	0.889 (0.863–0.915)	<0.0001	0.853 (0.827–0.879)	<0.0001	0.945 (0.885–1.010)	0.0942
Hospitalisation						
Non-hospitalised	1.0		1.0		1.0	
Hospitalised	17.438 (17.042–17.843)	<0.0001	82.699 (80.305–85.165)	<0.0001	146.073 (136.574–156.232)	<0.0001
ICU	0.698 (0.674–0.723)	<0.0001	0.788 (0.760–0.816)	<0.0001	0.739 (0.681–0.802)	<0.0001
ETI	9.976 (9.384–9.976)	<0.0001	8.272 (7.974–8.580)	<0.0001	7.751 (7.145–8.409)	<0.0001
Pneumonia status						
Without pneumonia	1.0		1.0		1.0	
With pneumonia	3.128 (3.074–3.182)	<0.0001	2.237 (2.200–2.275)	<0.0001	2.488 (2.406–2.573)	<0.0001
Comorbidity						
Not present	1.0		1.0			
Hypertension	1.176 (1.156–1.197)	<0.0001	1.131 (1.111–1.151)	<0.0001	1.051 (1.011–1.092)	0.0126
Obesity	1.292 (1.269–1.317)	<0.0001	1.335 (1.308–1.362)	<0.0001	1.020 (0.967–1.076)	0.4722
Diabetes	1.234 (1.213–1.256)	<0.0001	1.131 (1.111–1.152)	<0.0001	1.090 (1.048–1.133)	<0.0001
Cardiopathy	0.959 (0.922–0.996)	0.0319	0.911 (0.874–0.949)	<0.0001	0.799 (0.748–0.854)	<0.0001
COPD	1.261 (1.150–1.383)	<0.0001	0.992 (0.948–1.037)	0.7172	0.908 (0.844–0.976)	0.0093
Asthma	1.096 (1.051–1.143)	<0.0009	0.917 (0.865–0.972)	0.0036	0.852 (0.747–0.971)	0.0160
Immunosuppressed	1.273 (1.206–1.345)	<0.0001	1.347 (1.267–1.431)	<0.0001	1.224 (1.120–1.337)	<0.0001
CKD	2.086 (2.011–2.163)	<0.0001	1.763 (1.669–1.831)	<0.0001	1.450 (1.371–1.533)	<0.0001
Other complication	1.317 (1.269–1.368)	<0.0001	1.242 (1.194–1.291)	<0.0001	1.404 (1.314–1.501)	<0.0001

COPD, chronic obstructive pulmonary disease; CRD, chronic kidney disease; ETI, endotracheal intubation; ICU, intensive care unit.
Odds for ICU and ETI estimated from hospitalised data.
[a] $n = 1,387,777$.
[b] $n = 2,289,037$.
[c] $n = 2,918,646$.

found that in an unvaccinated population, in general, more severe cases were seen with Delta infections than with Omicron infections,[19] which would have increased levels of CFR.

Few studies have assessed the effects of COVID-19 vaccination in Mexico. A study with 312 health workers concluded that the BNT162b2 COVID-19 vaccine was 100% effective against severe illness; however, only 22 individuals were vaccinated in the trial,[20] so the results should be interpreted with caution. In addition, in a small study ($n = 53$), Galán-Huerta et al.[14] found that in vaccinated individuals, mainly with the Cansino vaccine, patients with complete vaccination were less likely to develop severe COVID-19 disease requiring hospitalisation compared with those who received incomplete immunisation. Moreover, a preprint study comparing 793,487 vaccinated individuals with 4,792,338 unvaccinated individuals from December 2020 to September 2021 suggests that vaccination can decrease hospitalisation and death for adults (aged ≥18 years) in Mexico.[21]

Increasing odds of mortality in middle-aged or older individuals and for hospitalised patients during 2021–2022

The odds of COVID-19 mortality in patients who were hospitalised increased from 2020 to 2022. This result can be explained by a relatively low incidence of cases that required hospitalisation from the total cases observed during 2021 and 2022; nonetheless, a higher proportion of these cases had an adverse outcome (96% and 94% for 2021 and 2022, respectively), increasing the odds of mortality in hospitalised patients. Distribution of cases between age groups suggests that a greater number of the patients in these groups required hospitalisation. The results, as previously discussed, might also suggest that these inpatients were not vaccinated.

The years 2021 and 2022 were characterised by the emergence of new SARS-CoV-2 variants (i.e. Delta and Omicron)[13] but also by the implementation of the national COVID-19 vaccination

programme. As observed in the CFR, most of the positive cases that did not require hospitalisation during 2021 and 2022 had higher probabilities of survival than non-hospitalised cases in 2020. Several studies support the effectiveness of vaccination protocols in Mexico and other countries.[9,16–24] For example, Mhawish et al.[24] indicated that most COVID-19 patients admitted to ICU in Saudi Arabia were non-immunised patients. Acuti Martelluci et al.[22] analysed 313,068 unvaccinated and 966,626 vaccinated residents in Italy and found that patients receiving two or three vaccine doses showed 80–90% lower risk of COVID-19 hospitalisation or death compared with unvaccinated patients. Similarly, Muthukrishnan et al.[23] examined the effectiveness of the COVISHIELD vaccine (ChAdOx1 nCoV-19) in 1168 patients in India, showing that fully vaccinated patients who required hospitalisation had a higher likelihood of survival than unvaccinated inpatients. One of the limitations of the present study is that vaccination status of patients was not included in the analysis because of the unavailability of information in the data sets. Nevertheless, although a significant increase in the odds of mortality was observed for hospitalised cases during 2021 and 2022, the CFR decreased during those years, suggesting that hospitalisation could have been related to unvaccinated patients or those with incomplete immunisation. This theory is supported by evidence highlighting the reduction in hospitalisation of vaccinated COVID-19 patients.[14,22,24]

Vaccine hesitancy is multifactorial, but a survey in Mexico indicated that middle-aged and older individuals were significantly more likely to refuse any COVID-19 vaccine, regardless of its effectiveness.[25] Similarly, another study in Mexico reported that a young adult population (18–34 years) was most likely to get vaccinated.[26] This could partially explain the increase in odds of mortality in middle-aged and older individuals observed in the present study following the implementation of the national COVID-19 vaccination campaign. Distrust in federal government recommendations also seems to play a significant role in COVID-19 vaccine refusal in the Mexican population,[25] but this is known to be a global problem.[27] Other factors for COVID-19 vaccine refusal in Mexico include a perception of adverse effects, conspiracy theories (e.g. the virus was created by the government) and anti-vaccine feedback from social media, friends and/or family.[26,28,29] Indeed, misinformation from social media has played a pivotal role in the development of vaccine hesitancy.[30,31]

Implications

The present investigation is an observational study based on retrospective analysis of a large data set (6,657,667 COVID-19-positive patients) of recorded information on COVID-19 cases. The findings suggest that the implementation of the COVID-19 vaccination programme had a noticeable positive effect, despite the significant increase in the number of positive cases during 2021 and 2022. The study also highlights the relevance for constant surveillance of COVID-19 vaccine effectiveness and the need to improve the information collected in free data sets provided by the government (e.g. background clinical information, vaccination status and adverse effects following vaccination). This information could provide confidence in the effectiveness of COVID-19 vaccination and improve vaccination acceptance.

An important limitation from previous reports and this study includes the lack of specific information available to determine clinical conditions; for example, disease severity or lack of specific information on comorbidity and clinical tests, which may reveal more evidence on the final cause of deaths. As revealed by Parra-Bracamonte et al.,[10] hospitalised patients might have more accurate and accumulative data regarding comorbidities leading to bias in their relationship to death, evidenced by the accumulated proportion of specific clinical conditions (i.e. pneumonia). Furthermore, interactions between comorbidities may bias disease outcomes, as suggested for CKD, where patients showed an increasingly very high risk for death, also in association with major comorbidities (diabetes, hypertension and obesity).[11] In addition, the size of the country, regional diversity and other patient habits are factors that need to be considered when analysing observed variations so that better management of the course of infections and strategies to improve prognosis in patients can be implemented.[10] This is particularly true in observed odds of mortality, where endotracheal intubation and non-hospitalised patients have a greater risk for death than those admitted to the ICU. This observation has been previously explained as a possible overburdening of the healthcare system capabilities, where patients are not receiving appropriate escalation of care because of limited resources;[10,32] this, although not properly documented, was evident, especially during the peaks of pandemic.

Finally, the differences in CFR among the studied years could also be explained by changes in the characteristics of SARS-CoV-2 variants, particularly during C2022, where the emergence of the Omicron variant was characterised by a rapid spread in the human population, lower fusogenicity and attenuated pathogenicity.[33]

Conclusions

The present COVID-19 analysis highlighted the lack of complementary information to fully understand and interpret the disease dynamic in Mexico, despite the availability of a large registered database. Regardless of this limitation, the results are based on a large data set and indirectly support the beneficial effects of implementing COVID-19 vaccination protocols, leading to a significant reduction of CFR from 10% during 2020 to less than 1% during 2022 in Mexico. However, an increase in the odds of mortality in hospitalised patients was seen during 2021 and 2022, which could be related to the unvaccinated proportion of the infected population following vaccination hesitancy. Some important comorbidities that were previously associated with increased odds of mortality became non-significant as the vaccination programme was implemented and a large proportion of the population were vaccinated.

Author statements

Acknowledgements

The authors thank the Secretariat of Health, Mexican Government, for free access to the database used in this study.

Ethical approval

None required.

Funding

None declared.

Competing interests

None declared.

References

1. Berlin I, Thomas D, Le Faou AL, Cornuz J. COVID-19 and smoking. *Nicotine Tob Res* 2020:1–3. https://doi.org/10.1093/ntr/ntaa059.

2. The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19)—China, 2020. *China CDC Weekly* 2020;**2**:113–22.
3. Li X, Xu S, Yu M, Wang K, Tao Y, Zhou Y, et al. Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan. *J Allergy Clin Immunol* 2020. https://doi.org/10.1016/j.jaci.2020.04.006.
4. WHO. *COVID-19 Weekly epidemiological update. Edition 87. 12 April 2022*. World Health Organization; 2020. 25 April 2022.
5. Secretaría de Salud. *Informe Técnico Diario COVID-19 MÉXICO. 17/04/2022*. Gobierno de México, https://www.gob.mx/salud/documentos/informe-tecnico-diario-covid19-2022, 2020. [Accessed 18 April 2022].
6. The World Bank. *Mexico Overview*. https://www.worldbank.org/en/country/mexico/overview, 2020. [Accessed 14 June 2020].
7. ENSANUT. *Encuesta Nacional de Salud y Nutrición, Presentación de Resultados*. Secretaría de Salud, Instituto Nacional de Salud Pública, Instituto Nacional de Estadística y Geografía; 2018. 42.
8. Ritchie H, Mathieu E, Rodés-Guirao L, Appel C, Giattino C, Ortiz-Ospina E, et al. *Coronavirus pandemic (COVID-19)*. Published online at OurWorldInData.org. Retrieved from: https://ourworldindata.org/coronavirus; 2020 [Online Resource].
9. Mathieu E, Ritchie H, Ortiz-Ospina E, Roser M, Hasell J, Appel C, et al. A global database of COVID-19 vaccinations. *Nat Human Behav* 2021;**5**(7):947–53. https://doi.org/10.1038/s41562-021-01122-8.
10. Parra-Bracamonte GM, Lopez-Villalobos N, Parra-Bracamonte FE. Clinical characteristics and risk factors for mortality of patients with COVID-19 in a large data set from Mexico. *Ann Epidemiol* 2020;**52**:93–8. https://doi.org/10.1016/j.annepidem.2020.08.005.
11. Parra-Bracamonte GM, Parra-Bracamonte FE, Lopez-Villalobos N, Lara-Rivera AL. Chronic kidney disease is a very significant comorbidity for high risk of death in patients with COVID-19 in Mexico. *Nephrology* 2021;**26**(3):248–51. https://doi.org/10.1111/nep.13827.
12. Lara Rivera AL, Parra Bracamonte GM, López Villalobos N. Tasas de letalidad y factores de riesgo por COVID-19 en México. *Rev Cubana Salud Publica* 2021;**47**(4).
13. De-La-Cruz-Hernández SI, Barrera-Badillo G. The Delta variant triggers the third wave of COVID-19 in Mexico. *Disaster Med Public Health Prep* 2022:1–3. https://doi.org/10.1017/dmp.2022.49.
14. Galán-Huerta KA, Flores-Treviño S, Salas-Treviño D, Bocanegra-Ibarias P, Rivas-Estilla AM, Pérez-Alba E, et al. Prevalence of SARS-CoV-2 variants of concern and variants of interest in COVID-19 breakthrough infections in a hospital in Monterrey, Mexico. *Viruses* 2022;**14**(1):154. https://doi.org/10.3390/v14010154.
15. Secretariat of Health (Secretaría de Salud)-Mexican Government. *Vacunación COVID*. 2020. http://vacunacovid.gob.mx/wordpress/. [Accessed 18 April 2022].
16. Murillo-Zamora E, Trujillo X, Huerta M, Ríos-Silva M, Baltazar-Rodríguez LM, Guzmán-Esquivel J, et al. Decreased risk of COVID-19 pneumonia in children and adolescents during the Delta variant emergence. *Public Health* 2022;**204**:9–11. https://doi.org/10.1016/j.puhe.2021.12.017.
17. Bernal JL, Andrews N, Gower C, Gallagher E, Simmons R, Thelwall S, et al. Effectiveness of Covid-19 vaccines against the B. 1.617. 2 (Delta) variant. *N Engl J Med* 2021;**385**:585–94. https://doi.org/10.1056/NEJMoa2108891.
18. Liang LL, Kuo HS, Ho HJ, Wu CY. COVID-19 vaccinations are associated with reduced fatality rates: evidence from cross-county quasi-experiments. *J Global Health* 2021;**11**. https://doi.org/10.7189/jogh.11.05019.
19. Li M, Liu Q, Wu D, Tang L, Wang X, Yan T, et al. Association of COVID-19 vaccination and clinical severity of patients infected with Delta or Omicron variants—China, May 21, 2021–February 28, 2022. *China CDC Weekly* 2022;**4**(14):293. https://doi.org/10.46234/ccdcw2022.074.
20. Murillo-Zamora E, Trujillo X, Huerta M, Riós-Silva M, Mendoza-Cano O. Effectiveness of BNT162b2 COVID-19 vaccine in preventing severe symptomatic infection among healthcare workers. *Medicina* 2021;**57**:746. https://doi.org/10.3390/medicina57080746.
21. Bello-Chavolla OY, Antonio-Villa NE, Valdés-Ferrer SI, Fermín-Martínez CA, Fernandez-Chirino L, Ramírez-García D, et al. Effectiveness of a nation-wide COVID-19 vaccination program in Mexico. *medRxiv* 2022. https://doi.org/10.1101/2022.04.04.22273330.
22. Acuti Martellucci C, Flacco ME, Soldato G, Di Martino G, Carota R, Caponetti A, et al. Effectiveness of COVID-19 vaccines in the general population of an Italian region before and during the Omicron wave. *Vaccines* 2022;**10**(5):662. https://doi.org/10.3390/vaccines10050662.
23. Muthukrishnan J, Vardhan V, Mangalesh S, Koley M, Shankar S, Yadav AK, et al. Vaccination status and COVID-19 related mortality: a hospital based cross sectional study. *Med J Armed Forces India* 2021;**77**:S278–82. https://doi.org/10.1016/j.mjafi.2021.06.034.
24. Mhawish H, Mady A, Alaklobi F, Aletreby W, Asad T, Alodat M, et al. Comparison of severity of immunized versus non-immunized COVID-19 patients admitted to ICU: a prospective observational study. *Ann Med Surg* 2021;**71**:102951. https://doi.org/10.1016/j.amsu.2021.102951.
25. Ramonfaur D, Hinojosa-González DE, Rodriguez-Gomez GP, Iruegas-Nuñez DA, Flores-Villalba E. COVID-19 vaccine hesitancy and acceptance in Mexico: a web-based nationwide survey. *Rev Panam Salud Pública* 2021;**45**. https://doi.org/10.26633/RPSP.2021.133.
26. Delgado-Gallegos JL, Padilla-Rivas GR, Zúñiga-Violante E, Avilés-Rodríguez G, Arellanos-Soto D, Gastelum-Arias LJ, et al. Determinants of COVID-19 vaccine hesitancy: a cross-sectional study on a Mexican population using an online questionnaire (COV-AHQ). *Front Public Health* 2021;**9**. https://doi.org/10.3389/fpubh.2021.728690.
27. Lazarus JV, Ratzan SC, Palayew A, Gostin LO, Larson HJ, Rabin K, et al. A global survey of potential acceptance of a COVID-19 vaccine. *Nat Med* 2021;**27**:225–8. https://doi.org/10.1038/s41591-020-1124-9.
28. Guaracha-Basáñez G, Contreras-Yáñez I, Álvarez-Hernández E, Román-Montes CM, Meza-López, Olguín G, et al. COVID-19 vaccine hesitancy among Mexican outpatients with rheumatic diseases. *Hum Vaccines Immunother* 2021;**17**:1–10. https://doi.org/10.1080/21645515.2021.2003649.
29. Larrondo UA, Peña FS, Morales GJ. Desinformación, vacunas y Covid-19. Análisis de la infodemia y la conversación digital en Twitter. *Rev Lat Comunicación Soc (RLCS)* 2021;**79**:1–18. https://doi.org/10.4185/RLCS-2021-1504.
30. Meneses-Navarro S, Bermúdez-Urbina FM, Hernández-Reyes NL, Santos-Hernández NG, De Fuentes-Vicente JA. Percepciones de riesgo al covid-19 entre jóvenes de comunidades indígenas y rurales del sureste de México. *Cienc Lat Rev Científica Multidiscip* 2021;**5**(6):13231–46. https://doi.org/10.37811/cl_rcm.v5i6.1318.
31. Pierri F, Perry BL, DeVerna MR, Yang KC, Flammini A, Menczer F, et al. Online misinformation is linked to early COVID-19 vaccination hesitancy and refusal. *Sci Rep* 2022;**12**:1–7. https://doi.org/10.1038/s41598-022-10070-w.
32. WHO. *Coronavirus disease (COVID-19) situation report e 88*. Geneva, Switzerland: World Health Organization; 2020. 18 September 2022.
33. Suzuki R, Yamasoba D, Kimura I, Wang L, Kishimoto M, Ito J, et al. Attenuated fusogenicity and pathogenicity of SARS-CoV-2 Omicron variant. *Nature* 2022;**603**:700–5. https://doi.org/10.1038/s41586-022-04462-1.

Public Health 215 (2023) 91–93

Contents lists available at ScienceDirect

Public Health

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

Short Communication

Dog-related deaths registered in England and Wales from 2001 to 2021

J.S.P. Tulloch [a, *], J.A. Oxley [a, b], R.M. Christley [a, c], C. Westgarth [a]

[a] Department of Livestock and One Health, Institute of Infection, Veterinary and Ecological Sciences (IVES), The University of Liverpool, CH64 7TE, UK
[b] PDSA, Telford, UK
[c] Dogs Trust, London, UK

ARTICLE INFO

Article history:
Received 26 October 2022
Received in revised form
5 December 2022
Accepted 11 December 2022
Available online 16 January 2023

Keywords:
Dog bites
Dog attacks
Death
Fatality
UK

ABSTRACT

Objective: This study aimed to describe the incidence and demographics of fatal dog bites or strikes, as defined in English and Welsh mortality data (2001–2021).
Study design: A descriptive analysis of the Office for National Statistics registered deaths data set.
Methods: Individuals whose cause of death was defined as 'bitten or struck by a dog' were identified. The average annual number of dog-related deaths and trends in incidence were calculated. Age and sex demographics of victims were described.
Results: In total, there were 69 registered deaths, a mean of 3.3 (95% confidence interval 0.3–6.3) dog-related deaths per year, and a mean annual incidence of 0.59 (95% confidence interval 0.06–1.11) deaths per 10 million population. There was no year-on-year change in incidence. Of victims, 59% were male, 10% were <5 years, and 30% were ≥75 years.
Conclusions: Dog-related deaths are rare in England and Wales and have not increased between 2001 and 2021. Further contextual information about the incidents is needed to be able to develop public health strategies and interventions.

Introduction

Between January and November 2022, the media have reported that at least nine people have been killed by dogs in England and Wales and have suggested that fatal attacks are increasing.[1] The incidence of dog bite–related hospital admissions in England more than doubled between 1998 and 2018, with over 8000 individuals being admitted annually and males more likely than females.[2] Over this period, the incidence tripled in adults to 15.0 admissions per 100,000 population in 2018, whilst that in children (defined as 14 years and under) remained stable and high (annual mean of 14.4).[2] Twenty-five percent of all admissions were children; in 2018, 1518 children were admitted compared with 6871 adults.[2] For females, there were peaks in childhood (5–9 years) and in middle age (45–49 years), whilst for males, there was a sole peak in childhood (10–14 years). National data regarding emergency departments are not available, but data from a single tertiary paediatric hospital reported stable levels of attendance for a dog bite between 2016 and 2019 (15 attendances per month), with no differences in attendance by sex; 33.4% of attendees were 7–12 years and 26.5% were 1–3 years.[3] It is unknown whether national dog-related deaths have increased in line with hospital admissions and whether the demographics of the victims are similar.

This study aimed to describe English and Welsh mortality data where the cause of death was registered as a dog bite or strike. We hypothesised that mortality incidence trends would be rising similar to hospital admissions data.

Methods

The Office for National Statistics data set 'Deaths registered in England and Wales – 21st century mortality' was explored.[4] These data are collated through the official certification and registration of deaths, a legal requirement since 1837, as such they have almost complete population coverage of deaths occurring in England and Wales.[5] The data describe the annual number of registered deaths stratified by age group, sex, year, geographical region (only from 2013 onwards) and the underlying cause of death. The cause of death is decided on by considering data from medical practitioners

* Corresponding author.
E-mail address: jtulloch@liverpool.ac.uk (J.S.P. Tulloch).

https://doi.org/10.1016/j.puhe.2022.12.005

and coroners alongside information provided once a death has been registered. These data solely record the primary cause of death; secondary or contributing causes are not included in this data set. The cause of death is coded using the World Health Organisation's International Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10).

Deaths were identified using the ICD-10 code group W54 (Bitten or struck by a dog). This was the same code group previously used to identify individuals admitted to English hospitals due to dog bites.[2] The annual average number and population incidence of dog-related deaths were calculated using Office for National Statistics population data as the denominator, and trends were described. Incidence was stratified by geographical region and plotted. The age and sex of the victims were described, and differences in sex-stratified incidence were examined using Chi-squared analysis. The ICD-10 codes used were analysed to identify any contextual information about the location of where the dog bites occurred.

All statistical and spatial analyses were carried out using R language (version 3.2.0; R Core Team 2015). The results were deemed statistically significant where $P < 0.05$.

Results

Between 2001 and 2021, 69 individual deaths were registered as being caused by a dog bite or strike. The annual number of deaths was approximately normally distributed, and the mean number of dog-related deaths per year was 3.29 (95% confidence interval [CI] 0.30–6.27).

The mean annual incidence was 0.59 (95% CI 0.06–1.11) per 10 million (Fig. 1). The highest incidence of dog-related deaths was in 2011, with 1.07 (95% CI 0.44–2.20) per 10 million population. There was no obvious trend in incidence.

Regional variation of dog-related deaths incidence was evident. The North-West of England had the highest average annual incidence of 1.36 deaths per 10 million population (95% CI 0.67–2.48), and the East of England had the lowest, 0.18 deaths per 10 million population (95% CI 0.02–0.83; Fig. 1).

Fifty-nine percent of victims ($n = 41$, 59.4%; 95% CI 46.9–71.1) were male (0.67 cases per 10 million population [95% CI 0.48–0.91]), compared with 0.44 cases per 10 million (95% CI 0.29–0.63) in females ($P = 0.08$). There appeared to be three peaks in incidence among males; 75–84 years (75–79 = 2.01, 80–84 = 2.36), 50–54 years (1.02), and <5 years (0.86), and two in females; 75–79 years (1.53), and <5 years (1.21; Fig. 1).

The majority of ICD-10 codes were not used, and from 2020 onwards, the summary code 'W54' was solely used. Only two codes specified a location of the dog attack ($n = 42$, 60.9% of cases). Of these, 81.0% (95% CI 65.9–91.4) occurred at home (W54.0) and 19.0% (95% CI 8.6–34.1) occurred on the street or highway (W54.8).

Discussion

The findings of this study do not support our initial hypothesis that dog-related deaths in England and Wales were increasing in number between and including 2001–2021, and instead, they appear to be at a stable, low rate; averaging three deaths per year. For context, over the same period, on average, one person per year died due to a lightning strike (ICD-10 code: X33),[4] three after falling from a tree (ICD-10 code: W14),[4] and 1548 in road accidents.[6] Data have yet to be officially collated for 2022, but it may record the greatest number of deaths for over 20 years. However, this must be treated with caution as one anomalous year does not make an increasing trend. Nevertheless, the media reports are concerning, and we recommend that data need to be reviewed annually to identify whether any trends of dog-related deaths emerge.

The calculated incidence of dog-related deaths (0.59 deaths per 10 million population) is slightly lower than in the United States of America (1.1 deaths per 10 million population) and similarly stable.[7] However, the annual European Union (EU) incidence, between 1995 and 2016, has been estimated as 0.9 per 10 million population, with an annual increase of 5.6%,[8] ranging from 0 in Ireland and Luxembourg to 4.5 in Hungary, with a median of 0.6 deaths per 10 million across the EU nations. It can be concluded that the incidence of dog-related deaths in England and Wales are not anomalous and is similar to many other North American and European

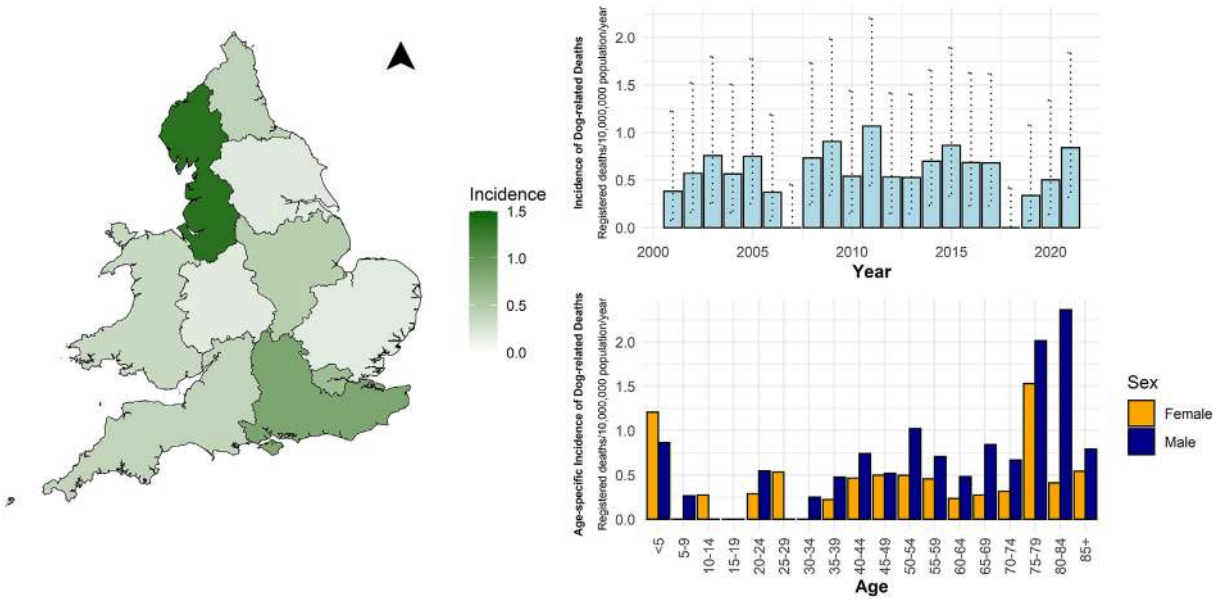


Fig. 1. Annual incidence annual incidence (cases per 10 million population per year) of dog-related deaths in England and Wales, stratified by region (2013–2021), year, and age groups (2001–2021).

countries. However, the trend data do not match the EU, which saw an increase above and beyond the growth in the dog population.[8] This is similar to English hospital admissions for dog bites, which saw an increasing rate of admissions that was greater than rate of growth in the dog population.[2]

Furthermore, across many North American and European nations, victims are predominantly male, with peak deaths in young children and the elderly.[8–10] Numerically, more males were seen here, but this was not statistically significant. Here, there was a high proportion of deaths seen in the <5 years (10.1%) and ≥75 years (30.4%).

There appears to be a disconnect between the stable incidence of dog-related deaths and the increasing incidence of hospital admissions for dog bites in England, which more than doubled over a similar period.[2] It is unknown why dog-related deaths are not increasing at the same rate. It may be that dog bites leading to minor injuries are increasing, but the most severe injuries (including fatal ones) are not increasing. Without further contextual and medical information about the nature of the bite or attack and the type and severity of injury, it is not possible to draw a conclusion.

The geographical spread of deaths was inconsistent with the spatial distribution of hospital admissions.[2] Hospital admissions data identified hotspots in areas that would have equated to the regions of the North West (similarly highest death incidence out of the 10 regions) and North East (only 5th out of the 10 regions), whilst the area with the lowest hospital admissions was London (3rd highest death incidence). This is most likely due to the low number of deaths in each region, resulting in the modifiable area unit problem form of bias. As such, these spatial patterns must be considered with a degree of caution.

The demographics of those killed vs. those admitted to hospital also differed slightly. In both cases, peaks were seen in young children but were younger in the registered deaths data set.[2] Peak adult admissions occurred in middle age (40–49 years), whilst peak adult deaths occurred in >75 years. This is likely due to the increased vulnerability of younger children and the elderly to fatal injury. The type of location of the dog-related incident, as recorded by ICD-10 codes, is comparable with 81.0% of individuals killed at home compared with 83.9% of those admitted to hospital, and 19% and 12% on the street, respectively.[2]

The major limitation of these data is the lack of contextual information about the events leading up to the victims' death. Understandably, as this is not the data set's primary purpose, there is no data concerning the dogs involved, the nature of the attack and events preceding it, and the extent and severity of injury. From a public health perspective, these data are critical so that interventions can be developed that may prevent these rare events from happening in the future. In addition, as no secondary or contributory causes of deaths are reported within these data, we do not know the degree our incidence figures are an over, or under, estimate. Due to the low number of deaths, when these data are stratified by demographics or geographical areas, group size becomes small and so only crude descriptive analysis can occur.

Conclusions

Dog-related deaths are very rare events within England and Wales and have not been increasing in incidence to 2021. The incidence and demographics are broadly similar to those of other European and North American nations. More contextual information is needed about dog-related deaths to be able to elicit change and develop effective public health strategies.

Author statements

Ethical approval

No ethical approval was needed, as this was an analysis of publicly available data, with no personally identifiable information. Patients were not involved in the design, nor participated in the research, of this study.

Funding

J.A.O.'s PhD studentship is funded by the Dogs Trust. J.T. and C.W. received no funding.

Competing interests

J.A.O. is employed as a scientific officer at the PDSA. C.W. has been a consultant for Forthglade Pet Food, Royal Canin and the Waltham Petcare Science Institute and has received financial renumeration. J.T. and R.C. have no conflicts of interest to declare.

Author contributions

J.T. conceived and designed the work, analysed the data and drafted the article. All authors interpreted the data, helped to revise the article and approved the final submitted article. All authors agree to be accountable for the accuracy and integrity of this work.

References

1. Defra Press Office. *Defra response to reported recent increase in fatal dog attacks*. 2022. https://deframedia.blog.gov.uk/2022/10/06/defra-response-to-reported-recent-increase-in-fatal-dog-attacks/. [Accessed 10 October 2022].
2. Tulloch JSP, Owczarczak-Garstecka SC, Fleming KM, Vivancos R, Westgarth C. English hospital episode data analysis (1998 – 2018) reveal that the rise in dog bite hospital admissions is driven by adult cases. *Sci Rep* 2021;**11**(1):1767. https://doi.org/10.1038/s41598-021-81527-7.
3. Tulloch JSP, Minford S, Vicky P, Rotheram M, Christley RM, Westgarth C. Paediatric Emergency Department dog bite attendance during the COVID-19 pandemic: an audit at a tertiary children's hospital. *BMJ Paediatr Open* 2021;**5**(1):e001040. https://doi.org/10.1136/bmjpo-2021-001040.
4. Office for National Statistics. *Deaths registered in England and Wales – 21st century mortality*. 2022. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/the21stcenturymortalityfilesdeathsdataset. [Accessed 20 September 2022].
5. Office for National Statistics. *Mortality statistics in England and Wales QMI*. 2022. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/methodologies/mortalitystatisticsinenglandandwalesqmi. [Accessed 20 September 2022].
6. Department for Transport. *Reported road accidents, vehicles and casualties tables for Great Britain*. 2022. https://www.gov.uk/government/statistical-data-sets/reported-road-accidents-vehicles-and-casualties-tables-for-great-britain. [Accessed 20 September 2022].
7. Centers for Disease Control and Prevention. *National Center for Health Statistics Mortality Data on CDC WONDER*. 2022. https://wonder.cdc.gov/Deaths-by-Underlying-Cause.html. [Accessed 20 September 2022].
8. Sarenbo S, Svensson PA. Bitten or struck by dog: a rising number of fatalities in Europe, 1995–2016. *Forensic Sci Int* 2021;**318**:110592. https://doi.org/10.1016/j.forsciint.2020.110592.
9. Forrester JA, Weiser TG, Forrester JD. An update on fatalities due to venomous and nonvenomous animals in the United States (2008–2015). *Wilderness Environ Med* 2018;**29**:36–44. https://doi.org/10.1016/j.wem.2017.10.004.
10. Raghavan M. Fatal dog attacks in Canada, 1990-2007. *Can Vet J* 2008;**49**:577–81.

Public Health 215 (2023) 75–82

Contents lists available at ScienceDirect

Public Health

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

Original Research

Economic evaluation of a school-based strategy to prevent overweight and obesity in French adolescents: insights from the PRALIMAP randomised trial

A.Y. Omorou [a, c], F. Manneville [a, c, *], H. Achit [a], J. Langlois [b], K. Legrand [a], E. Lecomte [b], S. Briançon [c]

[a] CHRU-Nancy, INSERM, Université de Lorraine, CIC, Epidémiologie Clinique, F-54000, Nancy, France
[b] National Conservatory of Arts and Crafts (CNAM), Nancy, France
[c] Université de Lorraine, APEMAC, F-54000, Nancy, France

ARTICLE INFO

Article history:
Received 21 July 2022
Received in revised form
31 October 2022
Accepted 30 November 2022
Available online 14 January 2023

Keywords:
Adolescents
Obesity
Prevention
Economic analysis

ABSTRACT

Objectives: This study aimed to provide cost-effectiveness and budget impact analyses of a school-based overweight/obesity screening and care prevention strategy among adolescents.
Study design: Cost-effectiveness and budget impact analyses.
Methods: Data from 3538 adolescents who participated in a school-based randomised controlled trial in the Northeast of France were used. Costs (from a public payer's perspective) included screening for overweight and obesity and subsequent care. Effectiveness was measured as the change in body mass index (kilogram per square metre), prevalence of overweight/obesity, moderate physical activity energy expenditure, duration and frequency and total sitting time. The incremental cost-effectiveness ratio was calculated, and a budget impact analysis was conducted.
Results: The screening and care strategy resulted in an incremental cost-effectiveness ratio of €1634.48 per averted case of overweight/obesity and €255.43 per body mass index unit decrease. The costs for increasing moderate physical activity by 1000 metabolic equivalent of task-min/week, duration by 60 min/week and frequency 1 day/week were €165.28, €39.21 and €93.66 per adolescent, respectively. Decreasing total sitting time by 60 min/week had a cost of €8.49 per adolescent. The cost of implementing the strategy nationally was estimated to be €50.1 million with a payback period from 3.6 to 7.3 years.
Conclusions: The screening and care strategy could be an efficient way to prevent overweight and obesity among adolescents. Future studies should investigate how the current results could be achieved in schools with different settings and thus justify its relevance for overweight and obesity prevention to policy-makers.

Introduction

The prevalence of overweight and obesity has rapidly increased over recent decades, and it has become a major public health issue because of the health consequences of these conditions.[1,2] The impact of overweight and obesity is not limited to population health; it also imposes a heavy economic burden on nations, resulting, in part, from the health expenditure generated by the treatment of overweight and obesity and related chronic conditions.[3] In support of this, the economic consequences of overweight and obesity in 52 countries from the Organisation for Economic Cooperation and Development, the European Union and the G20 were estimated to cost US$ 425 billion per year, accounting for 0.45%–1.62% of the countries' gross domestic products.[4] In France, the economic burden of overweight or obesity represents approximately 0.8% of the gross domestic product and accounts for nearly US$ 25 billion (i.e. 6%) of total health expenditure.[5]

Among existing measures to address public health obesity concerns, preventing obesity in adolescents is among the most

* Corresponding author. INSERM CIC-1433 Epidémiologie Clinique, CHRU de Nancy, rue du Morvan, 54 505 Vandœuvre-lès-Nancy, France. Tel.: +33(0)3 83 85 93 00.
E-mail address: f.manneville@chru-nancy.fr (F. Manneville).

https://doi.org/10.1016/j.puhe.2022.11.025
0033-3506/

important, given that overweight adolescents often remain overweight in adulthood.[6] A systematic review showed that 70%–80% of adolescents with obesity are concerned about the persistence of obesity in adulthood;[6] therefore, effective early intervention could reduce future morbidity. Accordingly, many studies have analysed the effectiveness and economic incentives for obesity interventions. In their recent systematic review, Zanganeh et al. included 39 studies with an economic evaluation of interventions for childhood and adolescent obesity.[7,8] These studies produced different results because the programmes evaluated were different in design; many studies demonstrated an economic benefit of such interventions,[9–14] but others concluded that the related costs would not be socially sustainable.[15–17] On this basis, it may be difficult to assess the cost-effectiveness and economic sustainability of adolescent obesity prevention programmes without conducting a specific economic evaluation. In addition, only five studies included in Zanganeh et al.'s systematic review concerned adolescents, suggesting a need for an economic study in this age group.[8] All existing overweight and obesity prevention programmes that have demonstrated to be effective among adolescents should be economically evaluated. Such evaluations will help and guide policy-makers and programme planners in their decisions to efficiently prevent overweight and obesity.

In France, The PRomotion de l'ALIMentation et de l'Activité Physique (PRALIMAP) trial was conducted to evaluate the effectiveness of a school-based overweight and obesity prevention strategy among adolescents.[18] Although, the analysis showed a positive effect of the prevention strategy (i.e. a decrease in the prevalence of overweight and obesity, an increase in physical activity [PA] and a decrease in total sitting time [ST]), no economic evaluation was conducted after the trial.[19] The present study aimed to provide cost-effectiveness and budget impact analyses (BIAs) of an adolescent overweight and obesity prevention strategy (the PRALIMAP trial) compared with no strategy in the French context.

Methods

The PRALIMAP trial

This study used data from the PRALIMAP trial, a randomised controlled study assessing the effectiveness of three intervention strategies for overweight and obesity prevention among

Table 1
Characteristics of adolescents who completed the PRALIMAP trial ($n = 3538$).

Characteristic	*n* (%)
Age (years ± SD)	15.6 (± 0.7)
Sex	
Boys	1499 (42.4)
Girls	2039 (57.6)
High school type	
Vocational	546 (15.4)
General/technical	2992 (84.6)
Home area	
Rural	1432 (40.5)
Urban	2106 (59.5)
Overweight or obesity	
No	2825 (79.8)
Yes	713 (20.2)
Screening and care strategy	
No	1851 (52.3)
Yes	1687 (47.7)

SD, standard deviation.
Data are numbers, unless indicated otherwise.

Table 2
Direct costs of collective sessions.

Cost	Session 1	Session 2	Session 3	Session 4	Session 5	Session 6	Session 7
Implicated professionals (amount paid/session)[a]	Physician (200 €) and dietician (78 €)	Dietician (78 €)	Dietician (78 €) and psychologist (90 €)	Sport educator (60 €)	Sports educator (60 €) and psychologist (90 €)	Psychologist (90 €)	Physician (200 €), dietician (78 €), psychologist (90 €) and Sports educator (60 €)
Amount paid session	278 €	78 €	168 €	60 €	150 €	90 €	428 €
Travel allowance (45 €/professional)	90 €	45 €	90 €	45 €	90 €	45 €	180 €
Cost per session organised[b]	368 €	123 €	258 €	105 €	240 €	135 €	608 €
Number of times the session was conducted	9	8	7	7	7	7	7
Total cost	3312 €	984 €	1806 €	735 €	1680 €	945 €	4256 €

[a] Based on 2021 costs.
[b] Cost per session organised was the sum of the amount paid to professionals and the travel allowance.

adolescents in 24 state-run high schools (i.e. the only eligibility criteria) in northeastern France over two academic years between 2006 and 2009.

The three health promotion strategies were 'educational' (i.e. lectures and group work on eating behaviour (EB) and PA), 'environmental' (i.e. increasing the availability of fruits, vegetables, water and PA) and 'screening and care' (see next section). These strategies were in line with the Ottawa charter, which provides a framework for health promotion actions using five means, of which three are particularly relevant for the prevention of overweight and obesity among adolescents in a school setting: develop personal skills (educational strategy), create a supportive environment (environmental strategy) and reorient health services (screening and care strategy).[20]

Outcomes were assessed at baseline (T0) and 2 years (T2). Each high school was assigned to receive none, one, two or all three strategies according to a 2 × 2 × 2 factorial cluster randomisation design (for each intervention strategy, 12 high schools received the intervention and 12 did not). High schools gave parents an information letter to obtain parental consent. If parents did not want their children to participate, they could inform high schools by a letter indicating their refusal. Adolescents were also given written and oral information and had the right to not participate.

The trial was approved by the French ethics committee Commission Nationale de l'Informatique et des Libertés (no. 906312) and the French data protection authority (no. 906312) and was registered at ClinicalTrials.gov (no. NCT00814554).

The study design, methods and rationale are described in detail elsewhere.[18] A total of 3538 adolescents (aged 14–18 years, 57.6% girls, 20.2% with overweight or obesity [see Table 1]) completed the PRALIMAP trial, and their data were analysed. Reporting of this study followed the 2022 Consolidated Health Economic Evaluation Reporting Standards[21] (see Supplementary Table S1).

Of all strategies, only the screening and care strategy was shown to be effective in reducing the prevalence of overweight and obesity,[19] increasing PA and decreasing ST.[22] Therefore, an economic evaluation of the screening and care strategy compared to no strategy (i.e. usual practice in schools) was conducted from a public payer's perspective (this is the most likely funding source to implement such an intervention). The periods assessed were the duration of the intervention (cost-effectiveness analysis) and lifetime (BIA). No health economic analysis was originally planned for the PRALIMAP trial, but this should be conducted in the future to contribute to scientific evidence.

Screening and care strategy

The screening and care strategy consisted of school nurses screening adolescents for overweight or obesity (i.e. measuring their weight, height and waist circumference) and proposing, if necessary, group care management. The eligibility criterion to receive care was weight excess corresponding to a body mass index (BMI) greater than the International Obesity Task Force[23] age- and sex-specific overweight thresholds. When eligible, adolescents were registered with a care programme that comprised seven scheduled collective 1.5-h sessions (i.e. group educational sessions) provided at or outside of each school. These sessions were centred around the themes of healthy eating and PA and were led by a multidisciplinary team (i.e. a physician, dietician, psychologist and sports educator) belonging to a health network specialising in overweight and obesity prevention. Twelve of the 24 high schools were randomly assigned to the screening and care strategy, with a total of 1687 adolescents (all were screened by school nurses). The other high schools did not receive the screening and care strategy, with a total of 1851 adolescents.

Effectiveness outcomes

Two anthropometric and four behavioural outcomes were considered in this study. Anthropometric outcomes comprised differences in changes in the prevalence of overweight and obesity and BMI between the screening and care and no screening and care groups from T0 to T2. Behavioural outcomes included differences in changes in moderate PA energy expenditure (metabolic equivalent of task [MET].min/week), duration (min/week) and frequency (day/week), and total ST (min/week) from T0 to T2 between the screening and care and no screening and care groups. PA and total ST were measured using the International Physical Activity Questionnaire.[24] The effectiveness outcomes of the screening and care strategy were adjusted for the other two strategies (i.e. educational and environmental).[19,22]

Measurement of costs

The costs of implementing the screening and care strategy were measured. There were no costs related to the no screening and care strategy because it consisted of the usual practice. All costs were obtained from structures of the PRALIMAP coordination committee (expense monitoring during the trial) and are presented in euros (€).

Costs of screening

Costs related to screening activity were estimated by multiplying the average time of a screening by the 2021 average hourly wage of school nurses in France. The average duration of screening time was estimated to be 10 min per adolescent, and the average wage was estimated to be €35 per hour (including staff and travel costs). These costs included all participants in the screening and care group and were calculated by taking the two measurement times (T0 and T2) into account.

Costs of care

Overweight and obesity management costs included all 2006 costs related to professional member coordination, training before conducting collective sessions and working time. A full-time project manager was recruited for an equivalent of 3 months to coordinate the setting of the intervention (i.e. organisation of professionals' training sessions, planning and coordination of collective sessions). All professionals received a half-day training session before the organisation of collective sessions. The overall costs of collective sessions were calculated by taking the amount each professional (i.e. physicians, dieticians, psychologists and sports educators) was paid, based on 2021 costs, by the programme into account (see Table 2).

Statistical analyses

All analyses were carried out using Microsoft Excel software 2016 and conducted according to an intention-to-treat principle.

Incremental cost and effectiveness

Cost-effectiveness was analysed by the incremental cost-effectiveness ratio (ICER) for each effectiveness outcome (i.e. the

$$ICER = \frac{screening\ and\ care\ strategy\ costs - no\ screening\ and\ care\ strategy\ costs}{screening\ and\ care\ strategy\ effectiveness - no\ screening\ and\ care\ strategy\ effectiveness}$$

difference in cost between the screening and care and no screening and care strategies, divided by their difference in the effectiveness outcome from T0 to T2) as follows:

The ICER represented the average incremental cost associated with one averted overweight and obesity case, a decrease in BMI of 1 kg/m^2, an increase in energy expenditure of 1000 MET-min/week energy expenditure (representing approximately 60 min of moderate PA [3 METs], 5 days a week), an increase in PA duration of 60 min/week, an increase in moderate PA frequency of a 1 day/week and a decrease in total ST of 60 min/week.

Budget impact analysis

In addition to the cost-effectiveness evaluation, a BIA was conducted (for overweight and obesity outcomes only) to compare costs that would result from fully implementing the screening and care strategy nationally to cost savings generated during adulthood by the same strategy (see Supplementary methods for more details). Given there is insufficient evidence of the long-term effect (i.e. into adulthood) of interventions to prevent overweight and obesity among adolescents,[25–27] two scenarios were modelled under two hypotheses ([1] the effect of the strategy would be maintained into adulthood and [2] the effect of the strategy would decrease by 25% into adulthood). BIA was conducted by discounting the annual adult overweight and obesity care cost at 0% and 1.5% over 60 years (which represents the mean life expectancy in France at 20 years), as recommended in France.[28]

Results

Screening and care costs

The total intervention cost for all participants was estimated at €47,400, and the average cost per adolescent was estimated to be €28.1 (see Table 3). Screening represented the highest cost component (41.6%), followed by project management (29.5%) and collective sessions (28.9%).

Intervention effectiveness

The screening and care strategy resulted in a 1.71% greater reduction (−2.27% and −0.56% in the screening and care and no screening and care high schools, respectively) in the prevalence of overweight and obesity ($P = 0.04$), corresponding to 61 averted cases[19] (see Table 4). BMI values changed more favourably in the 12 high schools that received screening and care (+0.64 ± 1.44) than BMI values in the high schools that did not receive screening and care (+0.72 ± 1.49), with a 0.11 kg/m^2 greater reduction (95% confidence interval [CI]: −0.21, −0.01; $P = 0.03$).[19] In the 12 high schools that received screening and care, increases in moderate PA energy expenditure, duration and frequency were 170.0 MET-min/week (95% CI: 50.0, 291.0; $P = 0.005$), 43.0 min/week (95% CI: 12.0, 73.0; $P = 0.005$) and 0.3 day/week (95% CI: 0.1, 0.6; $P = 0.04$) greater, respectively, than those in the 12 high schools that did not receive screening and care.[22] The schools that received screening and care had a greater reduction in total ST that was 198.6 min/week lower (95% CI: −313.2, −83.9; $P = 0.0006$) than in schools that did not receive screening and care.[22]

Cost-effectiveness analysis

The results showed that the cost of averting one case of overweight or obesity was estimated to be €1634.48 (see Table 4). The mean cost of decreasing BMI by 1 kg/m^2 was estimated at €255.43 per adolescent. The estimated costs of increasing moderate PA energy expenditure by 1000 MET-min/week, the duration by 60 min/week and the frequency by 1 day/week were €165.28, €39.21 and €93.66 per adolescent, respectively. Decreasing total ST by 60 min/week was estimated to cost €8.49 per adolescent.

Table 3
Direct costs of the screening and care strategy.

Costs items	Details	Costs (€)	
		Group (% of total cost)	Per participant
Costs of screening		19,682 (41.6)[a]	11.67[b]
Number of participants to receive screening at baseline and the end of the intervention ($n = 1687$)			
	Total duration[c] (h)	281.2	
	Average wage rate of school nurses (€/h)	35	
Costs of care			419.97[d]
Number of participants to receive care ($n = 66$)			
Project coordination costs		14,000 (29.5)	212.12
	Coordination (3 months of equivalent full time: 3 × 3000 €)	9000	
	Training of professionals	4000	
	Others	1000	
Costs of collective sessions		13,718 (28.9)	207.85
Total cost		47,400	28.1[e]

[a] The cost of screening was obtained by multiplying the total number of screening measurements (i.e. 3374 [1687 at baseline and 1687 at the end of the intervention]) by the total screening duration (i.e. 281.2 h) by the average wage rate of a school nurse (i.e. 35 €/h).
[b] Cost of screening per screened adolescent.
[c] The total duration was obtained by multiplying the total number of adolescents (i.e. 1687) by the average duration of the screening per adolescent (i.e. 10 min) and converted in hours (i.e. divided by 60).
[d] Cost of care per cared for adolescent.
[e] Cost of screening and care strategy per screened adolescent.

Table 4
Incremental effectiveness, cost and the incremental cost-effectiveness ratio of the screening and care strategy ($n = 3538$).

Effectiveness outcome	T0, %/mean ± SD	T2−T0, %/mean ± SD	Incremental effectiveness, %/β [95% CI]	*P*-value	Averted cases	Incremental cost (€)	ICER
Overweight and obesity prevalence (%)							
Screening and care group ($n = 1687$)	17.09%	−2.27%	−1.71%	0.04	29[a]	47,400	€1634.48 per averted case of overweight and obesity
No screening and care group ($n = 1851$)	19.90%	−0.56%	Reference			Reference	
Body mass index (kg/m^2)							
Screening and care group ($n = 1687$)	21.37 ± 3.20	0.64 ± 1.44	−0.11 [−0.21; −0.01]	0.03	Not applicable	47,400	€255.43 per kg/m^2 decrease per screened adolescent
No screening and care group ($n = 1851$)	21.69 ± 3.77	0.72 ± 1.49	Reference			Reference	
Moderate physical activity							
Energy expenditure (MET-min/week)							
Screening and care group ($n = 1687$)	587.8 ± 27.7	223.1 ± 37.9	170.0 [50.0; 291.0]	0.005	Not applicable	47,400	€165.28 per 1000 MET.min/week increase per screened adolescent
No screening and care group ($n = 1851$)	665.8 ± 28.8	74.3 ± 37.7	Reference			Reference	
Duration (min/week)							
Screening and care group ($n = 1687$)	147.0 ± 6.9	55.8 ± 9.5	43.0 [12.0; 73.0]	0.005	Not applicable	47,400	€39.21 per 60 min/week increase per screened adolescent
No screening and care group ($n = 1851$)	166.5 ± 7.2	18.6 ± 9.4	Reference			Reference	
Frequency (days/week)							
Screening and care group ($n = 1687$)	2.4 ± 0.1	0.5 ± 0.1	0.3 [0.1; 0.6]	0.04	Not applicable	47,400	€93.66 per 1 day/week increase per screened adolescent
No screening and care group ($n = 1851$)	2.5 ± 0.1	0.2 ± 0.1	Reference			Reference	
Total sitting time (min/week)							
Screening and care group ($n = 1687$)	2766.3 ± 26.1	15.4 ± 40.4	−198.6 [−313.2; −83.9]	0.0006	Not applicable	47,400	€8.49 per 60 min/week decrease per screened adolescent
No screening and care group ($n = 1851$)	2729.6 ± 23.6	209.3 ± 38.8	Reference			Reference	

β, regression coefficient; CI, confidence interval; ICER, incremental cost-effectiveness ratio; MET, metabolic equivalent of task; SD, standard deviation.
[a] The number of averted cases was calculated by multiplying the number of adolescents in the screening and care group by the incremental effectiveness (i.e. 1.71%).

Budget impact analysis

Nationally, the total number of eligible adolescents was estimated to be 1,782,172 (Table 5). Considering a cost of €28.1 per adolescent, the implementation cost of the screening and care strategy was estimated at €50.1 million at the national level. Furthermore, if the screening and care strategy was implemented, a 1.71% decrease in the prevalence of overweight and obesity would be obtained and would correspond to a total of 30,476 eligible adolescents. Considering the persistence rate of overweight and obesity from adolescence to adulthood to be 70%, a total of 21,334 cases of overweight or obesity in adulthood would be averted in scenario 1. These cases would result in an annual care cost savings for the public payer of €13.8 million. Scenario 2 would lead to a total of 16,000 averted cases of overweight and obesity in adulthood and would result in an annual care cost savings of €10.4 million. The payback periods were 3.6 and 4.8 years under scenarios 1 and 2, respectively. With a discount of 1.5%, the mean additional care cost was €430 per individual per year, which translated into payback periods of 5.5 and 7.3 years for scenarios 1 and 2, respectively.

Discussion

This study presented cost-effectiveness and BIAs of an overweight and obesity screening and care strategy among French school-aged adolescents from the public payer's perspective. The cost of an averted case of overweight or obesity, a one-unit decrease in BMI, a 1000 MET-min/week increase in moderate PA and a 60-min decrease in total ST were €1634.48, €255.43, €165.28 and €8.49, respectively. In the case of a national implementation of the screening and care strategy, the payback period was estimated to range from 3.6 to 7.3 years. In light of these findings, school-based interventions of this type are likely to be cost-effective (i.e. in reference to the annual overweight or obesity care cost of an adult) uses of public funds and warrant consideration by policy-makers and programme planners.

The literature suggests that implementation of the screening and care intervention in high schools has the potential to make a cost-effective contribution to the reduction in the prevalence of overweight and obesity, increase in PA and decrease in total ST during adolescence. A study based on the screening and management of obesity among 6- to 12-year-old children showed an intervention cost of $237 (2014 US dollars [€200]) per BMI unit reduced, which is in line with the current findings.[29] Furthermore, in their 2-year school-based PA intervention targeting adolescents, Sutherland et al. estimated the costs per adolescent to avert a one-unit BMI gain and to increase the duration of moderate-to-vigorous PA by 1 min at $1408 (2014 Australian dollars [€870]) and $56 (€35), respectively.[30] These costs are substantially greater than those estimated in the present study, potentially due to differences in the year of the considered costs, and may suggest a better cost-effective ratio in school-based interventions that include PA and eating behaviour components, such as the PRALIMAP trial.

In France, in 2019, the mean life expectancy at 20 years of age was estimated at 60.3 and 66.1 years for men and women, respectively.[31] Compared with these estimations, the durations for the payback period shown in this study are relatively small (from 3.6 to 7.3 years), even when considering a lower life expectancy (approximately 4–10 years less according to Lung et al.[32]) among individuals with overweight or obesity than those with healthy weight status.

Transferability

The PRALIMAP trial was conducted in four northeastern French departments. From the perspective of its implementation at the national level, the question of its transferability warrants discussion. Transferability refers to the extent to which the measured effectiveness of an applicable intervention could be achieved in another setting[33] and depends on the target population, implementation conditions, professionals and environment.[34] Northeastern France is characterised by a higher prevalence of overweight and obesity than some other regions.[35] If implemented in such regions, and given that screening cost did not depend on the prevalence of overweight and obesity, the cost-effectiveness ratio could increase. Previous work evidenced three important aspects to consider in the transferability of the PRALIMAP intervention: (1) a multidisciplinary approach (interdisciplinary teamwork and support by managers); (2) a participatory process (involvement of stakeholders in setting goals and allowing them to adapt the intervention if necessary); and (3) support for knowledge transfer

Table 5
Budget impact analysis of the screening and care strategy on overweight and obesity by scenario.

Input	Scenario 1: constant effect of the intervention in adulthood	Scenario 2: 25% decrease of effect of the intervention in adulthood
Target population[a]	1,782,172	1,782,172
Cost of the intervention per adolescent (€)	28.1	28.1
Estimated total cost of the intervention (€)[b]	50,079,034	50,079,034
Potential averted cases in adults[c]	21,334	16,000
Discount rate: 0%		
Estimated additional care cost per individual per year for overweight and obesity (€)[d]	648	648
Total estimated additional care cost per year (€)[e]	13,824,432	10,368,000
Payback period (years)[f]	3.6	4.8
Discount rate: 1.5%		
Estimated additional care cost per individual per year for overweight and obesity (€)[g]	430	430
Total estimated additional care cost per year (€)[e]	9,173,620	6,880,000
Payback period (years)[f]	5.5	7.3

[a] Estimated from the 'Direction de l'évaluation, de la prospective et de la performance' (French Ministry of National Education, Youth and Sports).
[b] Obtained by multiplying the total target population by the cost per adolescent.
[c] Obtained by multiplying the total target population by the screening and care strategy effectiveness (i.e. 1.71%) and by the persistence proportion of overweight and obesity from adolescence to adulthood (i.e. 70%).
[d] These costs were estimated from Emery C et al.[34] and referred to the direct medical costs of adults with overweight or obesity (consumption of care and medical goods presented for reimbursement).
[e] Obtained by multiplying the estimated additional care cost per individual per year by the total number of potential averted cases.
[f] Obtained by dividing the estimated total cost of the intervention by the total estimated additional care cost per year.
[g] This cost was obtained by applying a discount rate of 1.5% each year over 60 years to the estimation from Emery C et al.[34] It represented the mean discounted additional care cost (per individual per year for overweight and obesity).

(mutual learning between stakeholders and researchers).[36] Taking the new context and environment in which the intervention is implemented into account is also crucial. For example, there could be a low participation rate due to the intervention location[37] or existing alternative programmes/local public health policies that could interact with the effectiveness of the strategy. It should be stated that the high schools that participated in the PRALIMAP trial were state run (as are a large majority of French high schools), whose organisations and programmes are similar. Thus, the fact that the structures are not fundamentally different could favour the transferability of the intervention. Notably, the results of the PRALIMAP trial led to the implementation of the PRALIMAP-INES (INEgalités de Santé) trial, which includes the screening and care strategy.[38,39] It was first implemented in northeastern France and was then transferred to Guadeloupe, a French island with a high prevalence of overweight and obesity (intervention in progress).[40] These elements could provide confidence in the transferability of the PRALIMAP trial. Within the framework of the international transferability of the intervention, it would be interesting to investigate how the screening and care strategy could be implemented or adapted in other countries and what its economic impact would be.

Strengths and limitations

The results of this study should be interpreted with consideration of its strengths and limitations.

The main strengths of this study include the BIA that was used to complete the cost-effectiveness analysis and provide important data for decision-making.[42] While the cost-effectiveness analysis provides a direct interpretation of the health and economic impact of the screening and care strategy, the BIA provides additional information to decision-makers on the financial consequences of nationally implementing the strategy. Second, outcomes on the effectiveness of the screening and care strategy are based on results from a 2-year randomised controlled trial with a large sample size.[19,22] Third, the use of two anthropometric and four behavioural outcomes allowed this study to report the ICER from a number of perspectives (obesity, PA and ST) and will facilitate comparisons across studies.

In terms of limitations, first, the estimated annual costs of overweight and obesity are from a study published in 2007, and costs of medical care have increased since that time.[41] In addition, these costs were estimated from individuals who consumed fewer medical goods and services than excluded ones (selection bias), and data were obtained by self-reporting (measure bias). These could have led to underestimated costs of overweight and obesity, but the results of the BIA are therefore conservative. However, there are no more recent estimations of the costs of overweight and obesity in the French context. Second, two hypothetical scenarios on the effective maintenance of the strategy were tested in the BIA and could lead to the overestimation or underestimation of the current results. However, there is no clear evidence of the long-term effectiveness of interventions from adolescence to adulthood on which these scenarios could be based.[27] Third, no modern technologies, such as social media or online portals, were included in the screening and care strategy; however, these tools may be inappropriate for screening adolescents (i.e. measurements by nurses must be done in person). If the strategy was implemented in the present day, modern technologies could be used to care for adolescents after screening (e.g. remote educational sessions), which would reduce the costs. Fourth, the use of a self-reporting questionnaire to measure PA could have led to an overreporting of PA by adolescents. However, the use of an objective measure, such as an accelerometer for several thousand adolescents, would have been difficult to implement (more expensive, less convenient). In addition, the questionnaire used is reliable and validated, which can provide confidence in its use.[24]

Conclusions

In conclusion, this study highlighted that the screening and care school–based strategy was effective in reducing the prevalence of overweight and obesity, the total ST and increasing moderate PA, with a relatively low cost of €28.1 per adolescent over two academic years. The costs, per adolescent, for avoiding one case of overweight or obesity, increasing moderate PA by 1000 MET-min/week and decreasing total ST by 60 min were €1634.48, €165.28 and €8.49, respectively. The national implementation of the strategy would cost €50.1 million and, each year, would avoid approximately €6.9–€13.8 million of cost increase caused by the morbidity for people with overweight or obesity. The strategy would be profitable after 3.6–7.3 years. The screening and care strategy could be an efficient way to prevent overweight and obesity among adolescents. Future studies should investigate how the current results could be achieved in schools with different settings and thus justify its relevance for overweight and obesity prevention to policy-makers.

Author statements

Acknowledgements

In the name of the PRALIMAP trial group, the authors thank all those who led to the implementation of this project.

Ethical approval

The trial was approved by the French ethics committee Commission Nationale de l'Informatique et des Libertés (no. 906312) and the French data protection authority (no. 906312) and was registered at ClinicalTrials.gov (no. NCT00814554).

Funding

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

Competing interests

The authors have no competing interests to declare.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2022.11.025.

References

1. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2014;**384**:766–81.
2. Gurnani M, Birken C, Hamilton J. Childhood obesity: causes, consequences, and management. *Pediatr Clin N Am* 2015;**62**:821–40.
3. Tremmel M, Gerdtham U-G, Nilsson P, Saha S. Economic burden of obesity: a systematic literature review. *Int J Environ Res Public Health* 2017;**14**:435.
4. OECD. *The heavy burden of obesity: the economics of prevention*. OECD; 2019. https://doi.org/10.1787/67450d67-en.
5. Ministère de l'Economie et des Finances. *Obésité: quelles conséquences pour l'économie et comment les limiter?* Trésor-Eco; 2016. https://www.tresor.economie.gouv.fr/Articles/90846524-d27e-4d18-a4fe-e871c146beba/files/1f8ca101-0cdb-4ccb-95ec-0a01434e1f34.

6. Simmonds M, Llewellyn A, Owen CG, Woolacott N. Predicting adult obesity from childhood obesity: a systematic review and meta-analysis. *Obes Rev* 2016;**17**:95–107.
7. Zanganeh M, Adab P, Li B, Frew E. Protocol for a systematic review of methods and cost-effectiveness findings of economic evaluations of obesity prevention and/or treatment interventions in children and adolescents. *Syst Rev* 2018;**7**:54.
8. Zanganeh M, Adab P, Li B, Frew E. A systematic review of methods, study quality, and results of economic evaluation for childhood and adolescent obesity intervention. *Int J Environ Res Public Health* 2019;**16**:485.
9. Wang LY, Yang Q, Lowry R, Wechsler H. Economic analysis of a school-based obesity prevention program. *Obes Res* 2003;**11**:1313–24.
10. Brown HS, Pérez A, Li Y-P, Hoelscher DM, Kelder SH, Rivera R. The cost-effectiveness of a school-based overweight program. *Int J Behav Nutr Phys Act* 2007;**4**:47.
11. McAuley KA, Taylor RW, Farmer VL, Hansen P, Williams SM, Booker CS, et al. Economic evaluation of a community-based obesity prevention program in children: the APPLE project. *Obes Silver Spring* 2010;**18**:131–6.
12. Magnus A, Haby MM, Carter R, Swinburn B. The cost-effectiveness of removing television advertising of high-fat and/or high-sugar food and beverages to Australian children. *Int J Obes* 2005 2009;**33**:1094–102.
13. Gortmaker SL, Long MW, Resch SC, Ward ZJ, Cradock AL, Barrett JL, et al. Cost effectiveness of childhood obesity interventions: evidence and methods for CHOICES. *Am J Prev Med* 2015;**49**:102–11.
14. Xu H, Li Y, Du S, Zhang Q, Liu A, Sun J, et al. Cost–utility and cost–benefit analyses of school-based obesity prevention program. *BMC Public Health* 2020;**20**:1608.
15. Moodie M, Haby M, Galvin L, Swinburn B, Carter R. Cost-effectiveness of active transport for primary school children – walking School Bus program. *Int J Behav Nutr Phys Act* 2009;**6**:63.
16. Moodie ML, Carter RC, Swinburn BA, Haby MM. The cost-effectiveness of Australia's active after-school communities program. *Obes Silver Spring* 2010;**18**:1585–92.
17. Moodie M, Haby MM, Swinburn B, Carter R. Assessing cost-effectiveness in obesity: active transport program for primary school children – TravelSMART Schools Curriculum program. *J Phys Act Health* 2011;**8**:503–15.
18. Briançon S, Bonsergent E, Agrinier N, Tessier S, Legrand K, Lecomte E, et al. PRALIMAP: study protocol for a high school-based, factorial cluster randomised interventional trial of three overweight and obesity prevention strategies. *Trials* 2010;**11**:119.
19. Bonsergent E, Agrinier N, Thilly N, Tessier S, Legrand K, Lecomte E, et al. Overweight and obesity prevention for adolescents: a cluster randomized controlled trial in a school setting. *Am J Prev Med* 2013;**44**:30–9.
20. World Health Organization. *Milestones in health promotion. Statements from Global Conferences*. World Health Organization; 2009.
21. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs, Carswell C, et al. Consolidated health economic evaluation reporting standards (CHEERS) 2022 explanation and elaboration: a report of the ISPOR CHEERS II good practices task force. *Value Health J Int Soc Pharmacoeconomics Outcomes Res* 2022;**25**. https://doi.org/10.1016/j.jval.2021.10.008.
22. Omorou AY, Langlois J, Lecomte E, Vuillemin A, Briançon S. Adolescents' physical activity and sedentary behavior: a pathway in reducing overweight and obesity: the PRALIMAP 2-year cluster randomized controlled trial. *J Phys Act Health* 2015;**12**:628–35.
23. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000;**320**:1240–3.
24. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003;**35**:1381–95.
25. Brown T, Moore TH, Hooper L, Gao Y, Zayegh A, Ijaz S, et al. Interventions for preventing obesity in children. *Cochrane Database Syst Rev* 2019;**2019**. https://doi.org/10.1002/14651858.CD001871.pub4.
26. Kwasnicka D, Dombrowski SU, White M, Sniehotta F. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. *Health Psychol Rev* 2016;**10**:277.
27. Heijden LB van der, Feskens EJM, Janse AJ. Maintenance interventions for overweight or obesity in children: a systematic review and meta-analysis. *Obes Rev* 2018;**19**:798–809.
28. Choix méthodologiques pour l'évaluation économique à la HAS. Haute Aut. Santé. https://www.has-sante.fr/jcms/r_1499251/fr/choix-methodologiques-pour-l-evaluation-economique-a-la-has (accessed 17.12.21).
29. Sharifi M, Franz C, Horan CM, Giles CM, Long MW, Ward ZJ, et al. Cost-effectiveness of a clinical childhood obesity intervention. *Pediatrics* 2017;**140**:e20162998.
30. Sutherland R, Reeves P, Campbell E, Lubans DR, Morgan PJ, Nathan N, et al. Cost effectiveness of a multi-component school-based physical activity intervention targeting adolescents: the 'Physical Activity 4 Everyone' cluster randomized trial. *Int J Behav Nutr Phys Act* 2016;**13**:94.
31. Mortality – demographic balance sheet 2019. INSEE. https://www.insee.fr/en/statistiques/2555227?sommaire=2382613#consulter-sommaire (accessed 18.11.20).
32. Lung T, Jan S, Tan EJ, Killedar A, Hayes A. Impact of overweight, obesity and severe obesity on life expectancy of Australian adults. *Int J Obes* 2019;**43**:782–9.
33. Wang S, Moss JR, Hiller JE. Applicability and transferability of interventions in evidence-based public health. *Health Promot Int* 2006;**21**:76–83.
34. Cambon L, Minary L, Ridde V, Alla F. Transferability of interventions in health education: a review. *BMC Public Health* 2012;**12**:497.
35. Chardon O, Guignon N, Guthmann J-P, Fonteneau L, Delmas M-C. *La santé des adolescents scolarisés en classe troisième*. Études et Résultats. 2014.
36. Trompette J, Kivits J, Minary L, Cambon L, Alla F. Stakeholders' perceptions of transferability criteria for health promotion interventions: a case study. *BMC Public Health* 2014;**14**. https://doi.org/10.1186/1471-2458-14-1134.
37. Bonsergent E, Thilly N, Legrand K, Agrinier N, Tessier S, Lecomte E, et al. Process evaluation of a school-based overweight and obesity screening strategy in adolescents. *Glob Health Promot* 2013;**20**:76–82.
38. Legrand K, Lecomte E, Langlois J, Muller L, Saez L, Quinet M-H, et al. Reducing social inequalities in access to overweight and obesity care management for adolescents: the PRALIMAP-INÈS trial protocol and inclusion data analysis. *Contemp Clin Trials Commun* 2017;**7**:141–57.
39. Briançon S, Legrand K, Muller L, Langlois J, Saez L, Spitz E, et al. Effectiveness of a socially adapted intervention in reducing social inequalities in adolescence weight. The PRALIMAP-INÈS school-based mixed trial. *Int J Obes* 2020;**44**:895–907.
40. DGOS. *Expérimentations en cours*. Ministère Solidar. Santé; 2021. https://solidarites-sante.gouv.fr/systeme-de-sante-et-medico-social/parcours-des-patients-et-des-usagers/article-51-lfss-2018-innovations-organisationnelles-pour-la-transformation-du/article/experimentations-en-cours (accessed 21.07.21).
41. Emery C, Dinet J, Lafuma A, Sermet C, Khoshnood B, Fagnani F. Évaluation du coût associé à l'obésité en France. *Presse Med* 2007;**36**:832–40.
42. Yagudina RI, Kulikov AU, Serpik VG, Ugrekhelidze DT. Concept of combining cost-effectiveness analysis and budget impact analysis in health care decision-making. *Value Health Reg Issues* 2017;**13**:61–6.

ELSEVIER

PUBLIC HEALTH

www.elsevier.com/locate/puhe

Editorial Office

Natalia Camicia
Public Health Editorial Office,
RSPH, John Snow House,
59 Mansell St., London, E1 8AN,
Tel.: +44 (0) 207 265 7331
Fax: +44 (0) 207 265 7301
E-mail: publichealth@rsph.org.uk

Gabby Favell
Editorial Assistant,
Royal Society for *Public Health*,
John Snow House,
59 Mansell St., London, E1 8AN,
E-mail: gfavell@rsph.org.uk

RSPH
ROYAL SOCIETY FOR PUBLIC HEALTH
VISION, VOICE AND PRACTICE

Public Health 215 (2023) 83–90

Contents lists available at ScienceDirect

Public Health

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

Original Research

Evolution of social mood in Spain throughout the COVID-19 vaccination process: a machine learning approach to tweets analysis

A. Turón, A. Altuzarra, J.M. Moreno-Jiménez, J. Navarro*

Grupo Decisión Multicriterio Zaragoza (GDMZ), Dpt. Economía Aplicada, Facultad de Economía y Empresa, Universidad de Zaragoza, Gran Vía 2, 50003, Zaragoza, Spain

ARTICLE INFO

Article history:
Received 18 July 2022
Received in revised form
5 December 2022
Accepted 8 December 2022
Available online 14 December 2022

Keywords:
COVID-19 vaccination process
Sentiment analysis
Machine learning
Multivariate statistics
Tweets
Social mood

ABSTRACT

Objectives: This paper presents a new approach based on the combination of machine learning techniques, in particular, sentiment analysis using lexicons, and multivariate statistical methods to assess the evolution of social mood through the COVID-19 vaccination process in Spain.
Methods: Analysing 41,669 Spanish tweets posted between 27 February 2020 and 31 December 2021, different sentiments were assessed using a list of Spanish words and their associations with eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy and disgust) and three valences (neutral, negative and positive). How the different subjective emotions were distributed across the tweets was determined using several descriptive statistics; a trajectory plot representing the emotional valence vs narrative time was also included.
Results: The results achieved are highly illustrative of the social mood of citizens, registering the different emerging opinion clusters, gauging public states of mind via the collective valence, and detecting the prevalence of different emotions in the successive phases of the vaccination process.
Conclusions: The present combination in formal models of objective and subjective information would therefore provide a more accurate vision of social reality, in this case regarding the COVID-19 vaccination process in Spain, which will enable a more effective resolution of problems.

Introduction

The COVID-19 outbreak has been declared a pandemic by the World Health Organization because of its high rate of spread, severity and its frequent outcomes of severe pneumonia, respiratory failure and death.[1] Vaccination has become the main available public resource against the pandemic. However, the prejudices or sentiments of the general public and political leaders, as reflected in social media, are having a significant impact on the progression towards achieving vaccination targets.[1,2]

Social media such as Twitter, Facebook, YouTube and LinkedIn, with billions of users worldwide,[3] represent the preferred sites for sharing, almost instantly and very easily, thoughts, feelings and opinions on all kinds of events.[4] Twitter[5] is one of the most active platforms with approximately 290.5 million monthly active users worldwide in 2020 and was projected to keep increasing up to over 340 million users by 2024.[6] Every second around 6000 tweets on average are tweeted, which corresponds to more than 350,000 tweets sent per minute, 500 million tweets per day and around 200 billion tweets per year.[7]

Tweets are real-time messages with a maximum length of 280 characters at a time. They can be analysed based on *hashtags*, which refer to the symbol (#) in Twitter (for instance: #COVID19), containing a combination of the word *hash* from 'hash mark' and the word *tag*, that marks something belonging to a specific category. Hashtags make it easy to quickly find messages about a topic of interest as well as to collect all the sentiments and opinions of people in one place or country.[8–11]

One of the most promising methods for content analysis in social media is sentiment analysis.[12,13] It can be understood as a set of approaches, techniques and tools that extracts people's opinions, feelings and thoughts from users' text data by means of natural language processing methods.[14] Sentiment analysis through social media is growing rapidly within the international scientific community as a useful tool to understand people's opinions and attitudes on any important situation or phenomenon that affects public opinion.[11,15] For instance, natural disasters,[11] the Syrian refugee crisis,[4] the UK-EU referendum,[16] the impact of Brexit,[17] presidential

* Corresponding author.
E-mail address: jnavarrol@unizar.es (J. Navarro).

https://doi.org/10.1016/j.puhe.2022.12.003

or general elections in the United States,[18,19] Indonesia[20] and India,[21] the world cup soccer tournament,[22] extremism in social media,[23] 2019 EVALI outbreak[24] and the COVID-19 outbreak.[25,26]

This article presents a new approach based on the combination of machine learning techniques, in particular, sentiment analysis using lexicons, and multivariate statistical methods to assess the evolution of social mood through the COVID-19 vaccination process in Spain via tweet messages. Sentiment analysis, or opinion mining, will allow us to carry out the quantitative scrutiny of those tweets by extracting subjective information from the detection of eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy and disgust) and the assessment of polarity (valence), that is, the neutral, positive or negative connotation of the language used. Multivariate statistical methods, or data mining, will provide figures and graphics that can synthesise objective information and knowledge about the vaccination process; in particular, properties of social structures and the patterns of relationships among actors.

The proposed methodology has been applied to the analysis of 41,669 tweets from February 2020 to December 2021. It shows how the opinions expressed in social media can be analysed, so that the social mood of citizens can be detected, opinion groups and their leaders can be identified, and social support for government measures can be evaluated.[27–30] The present combination in formal models of objective and subjective information about the vaccination process provides a more accurate vision of reality, which will enable a more effective resolution of problems.

Vaccination process in Spain

The vaccination strategy in Spain was published on 2 December 2020, with 11 updates up to the end of the considered period for analysis.[31] Four phases were defined according to available doses (see Table 1). The population groups to be vaccinated were established in order of priority , following an assessment based on criteria that incorporated the risk of exposure and transmission, the existence of previous serious illness, and the socio-economic impact of the pandemic on each population group.[32]

Methods

The methodological approach was based on Social Web Mining complemented with natural language processing and social network analysis. Messages were collected from social networks, preprocessed, and then their features were extracted to perform an analysis of society's opinion and mood regarding that critical event, and the way people related to each other and exchanged information on that event on social networks. The chart in Fig. 1 shows the methodological procedure that consists of three steps and three stages for each step.

Step 1: Corpus Determination

Stage 1.1. Data collection

We used a data set of 300,286 tweets in Spanish, posted between 27 February 2020 and 31 December 2021, that is, from the beginning of the pandemic until the end of the main stage of the vaccination process in Spain. The tweets were extracted from Twitter using the *twitterR* package, written in R programming language, accessing Twitter API 2.0. and searching in the full historical Twitter database. The search key was built from the following hashtags: #covid; #covid19; #*Yomevacuno* (I'm getting vaccinated); #*Yonomevacuno* (I'm not getting vaccinated); #*Negacionista* (denialist). The key string used to query the database was (*covid* OR *covid19*) AND (*Yomevacuno* OR *Yonomevacuno* OR *negacionista*).

It was referring to COVID and vaccination and to the pro- and anti-vaccine positions. The search terms were written in Spanish, and the condition that the messages be written in Spanish was added.

The attributes extracted from each tweet and its author were stored in two separate tables in the database according to the scheme shown in Table 2.

Other R packages such as *httr*, *RCurl* or *jsonlite* were used to extract the information from the Twitter API, in addition to *RMySQL* to manage the data through a MySQL database.

Stage 1.2. Data preprocessing

The tweets were preprocessed to eliminate all elements of the data that are susceptible to inconsistency or ambiguity, or, for reasons of efficiency, unnecessary in the subsequent analysis (punctuation marks, symbols or numbers, and words that do not provide meaning). This means that from a total of 7,377,533 words, 5,813,263 were preserved after the depuration; in other words, 21.20% of the words were suppressed. The preprocessing was carried out using the *stringr* R package.

Table 1
Spanish vaccination phases according to available doses.

Phase/description	Duration	Population group
Phase 0/Development, authorisation and evaluation	From February 27 till 18 December 2020 (1st update)	
Phase 1/First available doses	From 19 December 2020 till 26 February 2021 (4th update)	• Residents and staff in nursing homes and care centres for the elderly and the highly dependent • Front-line health and social personnel • Other health and social care staff • Non-institutionalised major dependents
Phase 2/More available doses	From 27 February 2021 till 11 May 2021 (7th update)	• Over 80 years • People between 70 and 79 and people with very high-risk conditions • People between 60 and 65 • People between 66 and 69 • Other health and social care workers • Workers with an essential social function • People between 50 and 59
Phase 3/Widely available vaccine	From 12th May 2021 till 31st December 2021	• People between 40 and 49 • People between 30 and 39 • People between 20 and 29 • People between 12 and 19 • People between 5 and 11 • Booster doses

Table 2
Structure of the database.

Tweet		Author	
Tweet ID	Text	Author ID	Registration date
Author ID	Hashtags	Author name	Location
Creation date	Is retweeted	Username	Description

Table 3
Filters for corpus determination.

Filter	Number of tweets
Tweets collected	300,286
Tweets containing location	188,392
Authors geolocated in Spain	28,285
Authors geolocated in Spain with indication of region	24,394
Tweets posted by authors geolocated in Spain	41,669

Stage 1.3. Geolocation of the authors

To select the tweets written by Spanish authors, the geographical location of the authors was identified, when possible, from the information contained in the location field. This was done by calling the *Nominatim* geocoding service, an Open Data project/of *OpenStreetMap*.[33] A total of 188,392 tweets were posted by authors that contained information in this field, of which Nominatim obtained a location determined by its latitude, longitude and country. It was shown that 28,285 authors were from Spain and writing in Spanish, of which 24,394 had indication of the region.

The study considered the tweets sent by these 28,285 Spanish authors. In total, there were 41,669 tweets that constituted the corpus of the study, being some of them retweets of other authors (Table 3).

Step 2: Social mood evolution

Stage 2.1: Social network analysis

The most relevant network interaction was considered to be the retweet because the number of retweets was very abundant in the corpus and the action of sharing or retweeting a text implied personal interest from the person who retweeted. Given the list of 28,285 Spanish users, all their messages that were retweets were selected, and the authors of the original message were extracted

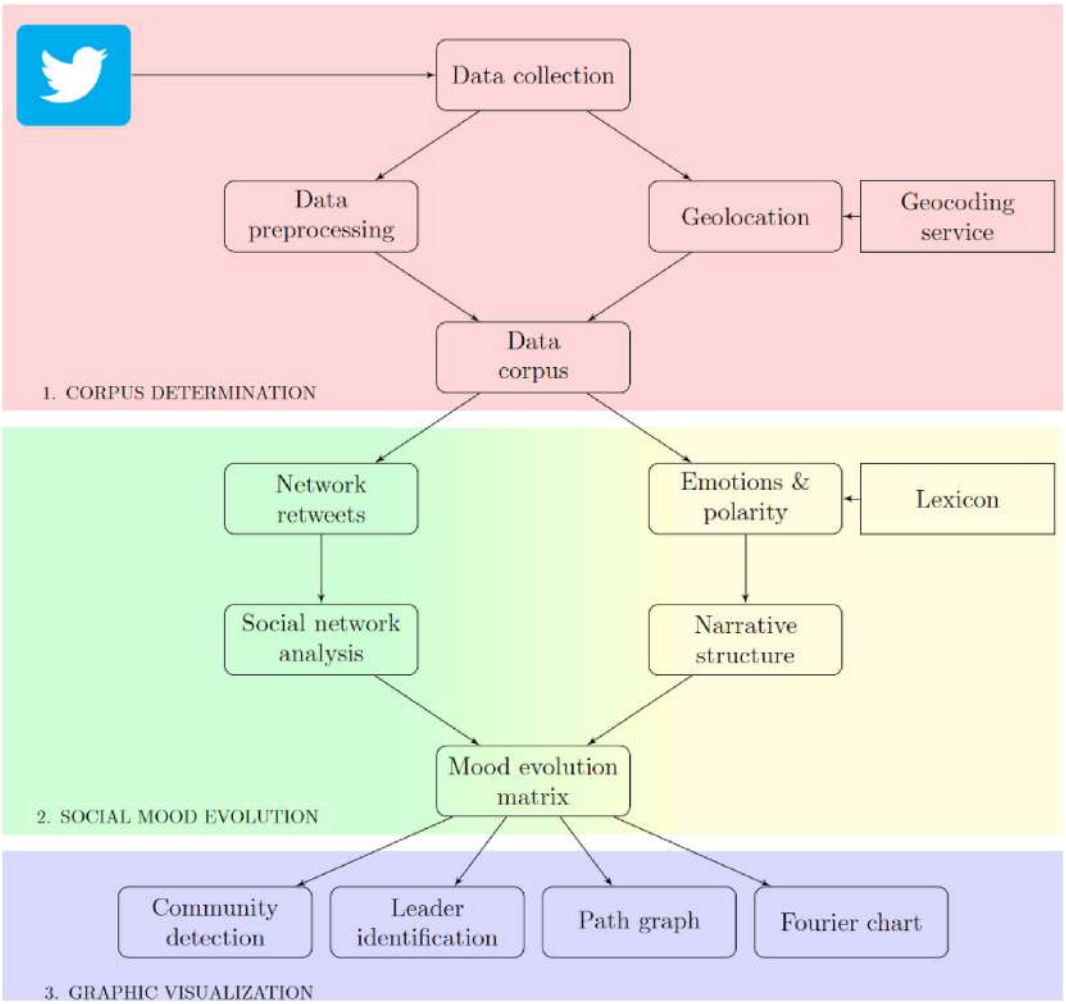


Fig. 1. Methodology flow diagram for the study of social mood evolution.

(although these may not be geolocated in Spain). A network was created based on the following methodological considerations:

- The network was a directed graph, the origin of each arc was the node corresponding to the author who retweeted a message and destination was the node that represented the author of the original tweet.
- The nodes were the users who had published tweets and retweets.
- The size of the nodes was proportional to the in-degree, representing the volume of retweets that has been made of their tweets.
- The colours of the nodes indicate communities. These communities have been calculated with the *Gephi software*,[34] which uses the algorithm described in.[35]
- The colour of the edges is the same as in the origin node, whereas their size is proportional to the number of messages from the destination node that the origin node has retweeted.
- The position of each node in the graph has been calculated using the *Force Atlas 2* algorithm,[36] an energy model for network spatialisation so that the more retweets a node has, the more focused it will be with respect to the nodes connected to it.

The resulting network contained 10,021 nodes and 17,340 edges, which represents a very low density, practically zero. Also, the average degree of the network is 1.73. This means that few retweets were made, and usually, the same authors were retweeted.

The analysis reveals the most influential users because of the size of their node (number of times a message of theirs has been retweeted) and their position within the cluster to which they belong (the more focused, the larger this size is). And the more compact a community is, the more relationships appear between its members. On the other hand, the different communities are closer to each other depending on how many nodes of each one are related to the other. The more relationships there are between two communities, the closer they would be.

Stage 2.2: Sentiment analysis

The 41,669 tweets were analysed, applying text mining by means of the *Syuzhet* 1.0.6 package[37] and *RStudio* 1.1.419, according to the general procedure already shown in Fig. 1.

As a first step, the sentiment was evaluated with *NRC Word-Emotion Association Lexicon Version 0.92*.[38–40] This lexicon is a list of English words and their associations with eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy and disgust) and two sentiments (negative and positive). For each tweet, the valence was also obtained, that is, the difference between the number of positive and negative words, as well as the number of words associated with each of the above emotions and sentiments. We then examined how emotions were distributed throughout the text. To do this, several descriptive statistics were obtained (minimum, maximum, Q1, Q3, mean, and median) with which an overall assessment of each tweet could be achieved.

Stage 2.3: Mood evolution Matrix

After performing the social network and sentiment analysis (Stage 2.2 and Stage 2.3), the result is a matrix where the rows are the different tweets (41,669) and the columns (40) are grouped into the following information blocks:

- Tweet variables (8 columns): id, author_id, date, text, clean text, hashtag, retweeted (yes or no), retweeted_id.
- User variables (14 columns): name, username, created_at, location, description, type, lat, lng, country, city, region, postal code, cod_region, id_region.
- Emotions (8 columns): eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy and disgust).
- Sentiments (4 columns): polarity (negative or positive), valence and number of sentiment words.
- Statistics (6 columns): six descriptive statistics (min, max, Q1, Q3, mean and median).

Results

This section presents the results corresponding to Step 3 of the methodology (Graphic Visualization). It includes illustrations of community detection, leader identification and path and Fourier graphs.

Community detection

Fig. 2 analyses the evolution of the retweet network during the phases of the process.

The most striking result is that two differentiated nuclei emerged, with very few interconnections between them, are distinguished in each phase: on the left, groups linked to the official sources of the Government and the health administrations of Spain, journalists and media (provaccine messages); on the right, accounts disseminating denialist and antivaccine messages. In Phase 0, there were 3818 users (2746 pro- and 1072 anti-vaccination); in Phase 1, 7758 users (5726 pro- and 2032 anti-vaccination); in Phase 2, 3510 users (2883 pro- and 627 anti-vaccination); and in Phase 3, 5637 users (2698 pro- and 2939 anti-vaccination). The composition and size of both pro- and anti-vaccine groups are clearly related to the variations produced in the social mood that will appear later in Fig. 3.

Leader identification

As can be seen in Table 4, there were several leaders involved in the different communities.

To better identify the leaders of the different communities, @sanidadgob corresponds to the official account of the Spanish Ministry of Health; @We_T_Resistance is an account positioned against the vaccination process; @salvadorilla (at the time Minister of Health of Spain); @rimbaudarth is an account positioned with the thesis of @We_T_Resistance; @publico_es is a media positioned in favour of the process; @Javier_CB is a very heterogeneous community with media presence but with very low activity on the network; and @daandina is a facultative working in public health. Clearly, the two most prominent leaders are the Government (1581 retweets) and the deniers (992 retweets).

Path and fourier graphs

The protocol described in Sections 3 and 4 (and Fig. 1) was applied to the 41,669 tweets. Fig. 3 shows the Fourier plot trajectory that represents emotional valence vs percentage of tweets (tweets date). From this analysis of tweets, we can see how the mental state or social mood of Spanish people has been changing through the different phases of the vaccination process (in different colours).

As shown in Fig. 3, the highest value of valence is found at Phase 1 (orange), between 4 and 6 January 2021, corresponding with the start of vaccination in Spain with Pfizer-BioNTech COVID-19 vaccine and the approval of Moderna COVID-19 (MD) vaccine by the European Medicines Agency. While the lowest value of valence is found at Phase 3 (green), between 4 and 6 August 2021, corresponding with the announcement of the need for booster doses and the debate on compulsory vaccination. On the other hand, we should note that the biggest fluctuations were produced in Phase 2

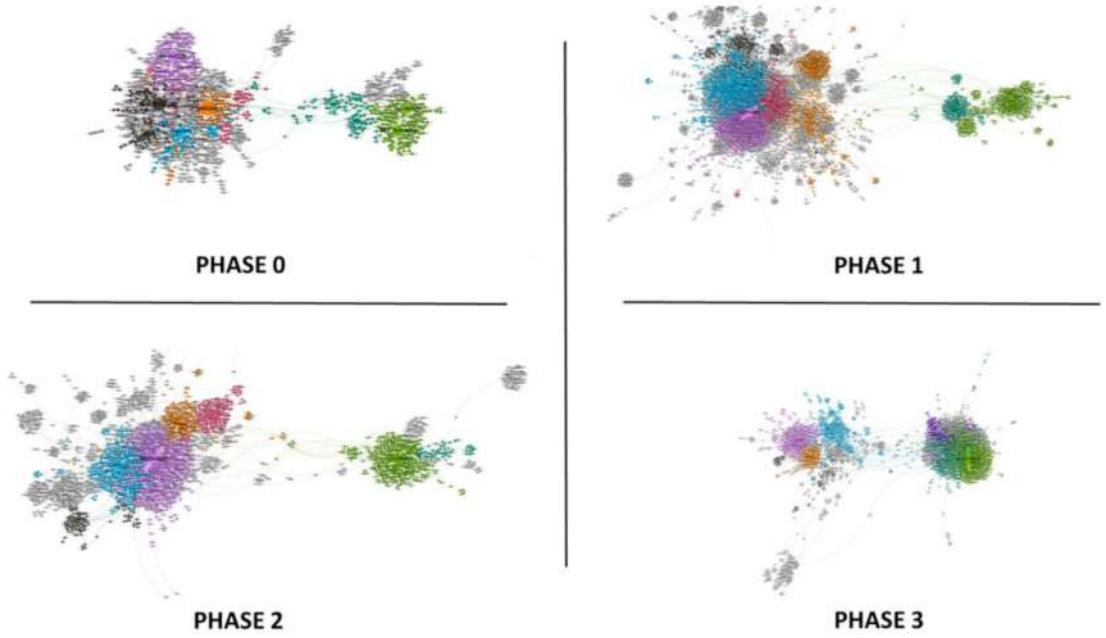


Fig. 2. Retweets network of the vaccination phases. The nodes are the users, and the arcs point goes from the retweeter to the author of the original tweet. The most retweeted authors are highlighted, and seven relatively clear clusters can be distinguished (each of them is formed by more than 2.5% of the total nodes and coloured in different colours). Within each cluster, those with highest number of retweets have been distinguished, appearing as the largest nodes in the graph.

(yellow) and Phase 3 (green) because of discordant health decisions on the Astra Zeneca vaccine.

Fig. 4 shows the percentage of words for each emotion according to each of the phases. It shows that the highest values for the main two emotions of the population during COVID-19 (fear and sadness)[41] were found at Phases 0 and 3. However, the highest value of joy and trust (more positive emotions) were shown in Phases 1 and 2, coinciding with the results obtained in Fig. 3 where the positive valences were in Phases 1 and 2.

The same pattern can be observed in Fig. 5 where we analyse the percentage of words for each phase according to each of the emotions. It is worth noting that the highest percentages of words

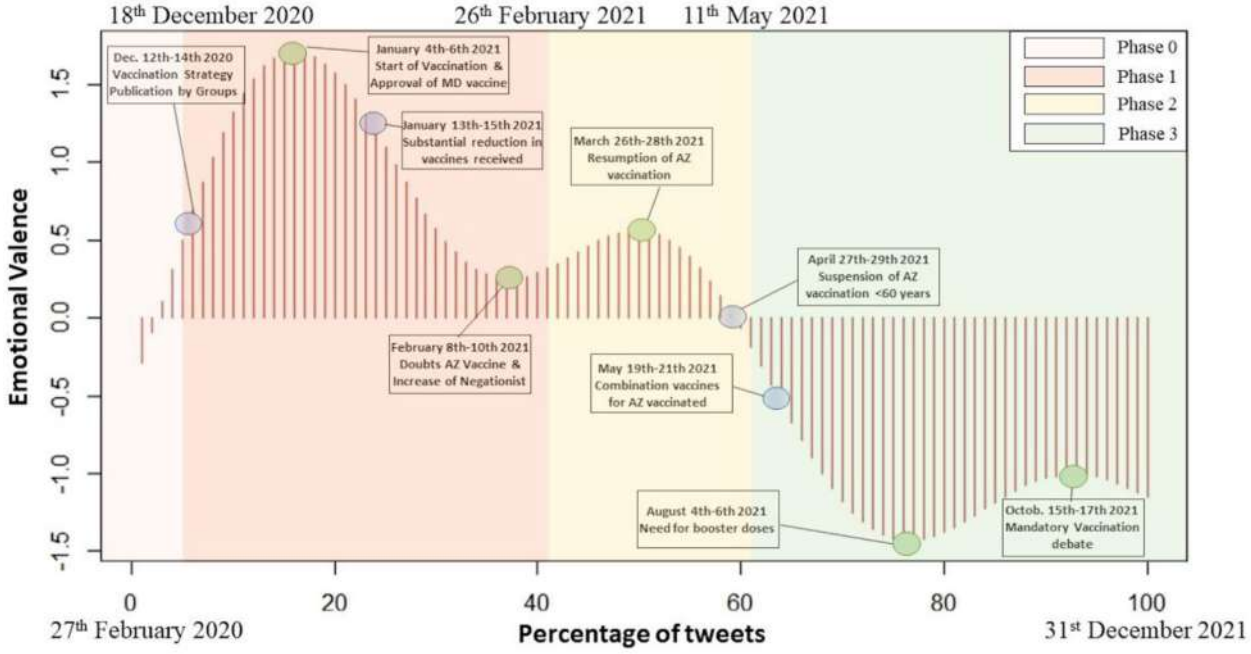


Fig. 3. Fourier plot trajectory of the tweets with the four phases (differently coloured). It represents emotional valence vs percentage of tweets (tweets date). In the upper side, the positive sentiments, and in the lower side, the negative ones. Local hotspots (green circles) and areas of trend change (purple circles) were marked by analysing the content of these tweets and relating them to relevant news and political decisions.

Table 4
Most retweeted authors in each community.

Community	Number of members (%)	Username	Number of retweets	Number of retweets (community)
Pink	1508 (15.05%)	@sanidadgob	1581	45.76%
Orange	446 (4.45%)	@daandina	42	6.87%
Black	774 (7.725)	@salvadorilla	347	17.08%
Fuchsia	426 (4.25%)	@Javier__CB	85	13.78%
Blue	1162 (11.60%)	@publico_es	158	7.52%
Green	1811 (18.07%)	@We_T_Resistance	922	23.28%
Emerald	292 (2.91%)	@rimbaudarth	171	34.76%

expressing the most negative emotions (anger, disgust, fear and sadness) are found in Phase 3, where the vaccines were widely available, but nevertheless, many doubts arose about the vaccination process with the news of the need for new doses or even compulsory vaccination. On the other hand, the most positive emotions (trust and joy) were in Phase 1, coinciding with the first available doses and the start of the vaccination process in Spain.

Discussion

This study has obtained a series of congruent results regarding the social networks involved, the evolution of social mood coupled with the dynamics of these networks, and the sentiment analysis represented in the plot trajectory. This overall congruence between the different kinds of obtained results may be interpreted as a very promising aspect of the approach.

Let us first point out that, regarding the evolution of social networks depicted in Fig. 2, the clustering dynamics during the four phases distinguished is surprisingly accurate, capturing the evolution of public opinion during the vaccination process. The analysis of the network of retweets not only shows the interconnections and clustering of the community of tweeters around interest groups but also shows how the structure of these groups varies throughout the process. It can be seen how public health decisions and other environmental circumstances that cause the changes in mood are translated not only into how tweeters are grouped but also who their referents are when it comes to sharing information. In addition, we can see in the network dynamics that clustering around two compact groups, of pro-vaccines and anti-vaccines, polarises the position of individuals in two communities with extremely few interconnections. These 'radical' divisions occur because of, and are exacerbated by, increasing conflict in communications about contentious topics such as lockdowns and compulsory vaccination.

Table 4 indicates the importance of public health communication from official sources (@sanidadgob and @salvadorilla) because their retweets from other users can reach far more people that are not following the official accounts. This means a cost-effective communication strategy for public health promotion.[42] In this regard, we may realise that most international political leaders are progressively turning to social networks to broadcast information about the pandemics, response plans, public health measures and connection with citizens.[43] This implies a series of strategic choices to use a more positive frame to influence opinion and action and to encourage compliance with public health norms and standards. The choice of positive frames may guide the national conversation away from seeking 'blame' for the pandemic towards a supportive mood necessary to implement the public health strategies required.[44] Finally, identifying and monitoring those social leaders whose opinions most closely reflect the needs or demands of society will contribute to make more realistic and effective public health decisions.

The prevalence of the different emotions during each of the phases shown in Figs. 4 and 5 would correlate well with the above. The high levels of anger, disgust, fear and sadness in phase 3 would document, as already said, the news about the new doses needed and the compulsory vaccination. The mental fatigue after the prolonged lockdowns and the stress for such long periods of uncertainty and pandemic fears are indeed reflected in the emotional arousal seen in these final phases.

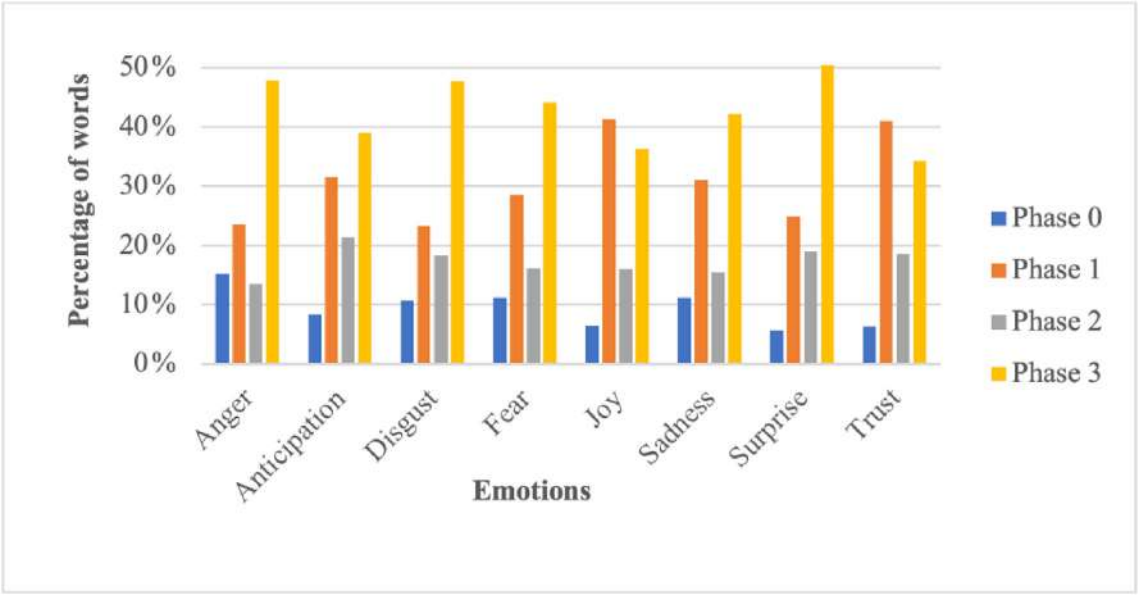


Fig. 4. Percentage of words per emotion according to each of the phases.

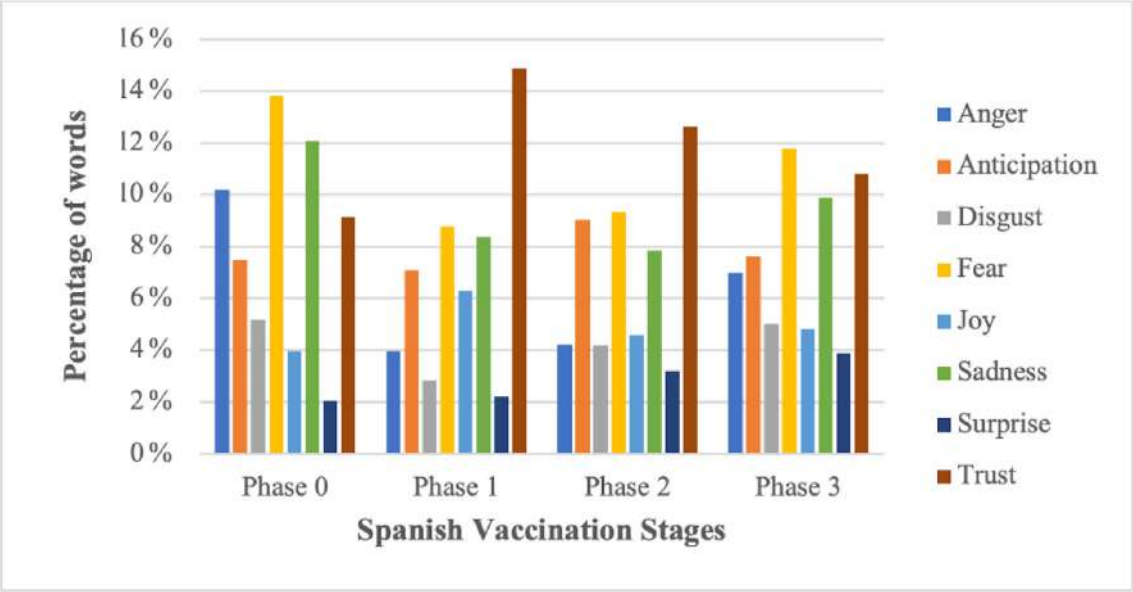


Fig. 5. Percentage of words per phase according to each of the emotions.

The specific results of sentiment analysis in the Fourier plot also show a remarkable congruence with the development of the four phases and the most notable events during the vaccination process. Although the way to obtain the valence of each tweet may look rather coarse, there is a considerable degree of theoretical sophistication in this evaluation of emotional valence. Some of the most accepted theories of emotions rely on two-dimensional spaces where valence becomes one of the fundamental dimensions.[45–49] The six basic emotions due to Paul Eckman[50] are generally maintained, although it is also generally accepted the need to enlarge these basic emotions.[51,52]

Sentiment analysis indeed offers an exciting panorama of emerging tools and paradigms to explain the emergence of social moods and emotional contagion phenomena that are so important in our societies, including the current 'epidemic of loneliness'.[53,54]

Looking at the limitations of the present approach, we have to consider the existing complementarity between the sentiment analysis technic using lexicons, as herein developed, and the machine learning and deep learning models (supervised and unsupervised).[55] Lexicon-based models are to be preferred where the data sets are small and the available computational resources limited under the condition of slightly lower performance.[56] The supervised models perform fine for the specific domain they have been trained. But this specific training becomes an important limitation for addressing different domains or brand-new topics such as the present COVID-19 pandemic. The unsupervised learning approaches do not hinge on the domain or topic of the training data, overcoming the difficulty of labelled training data collection and creation, although they need an extensive learning process and the subsequent computational resources. The hybrid technique is the combination of both lexicon and deep learning approaches. This combination improves the performance of classification, makes the detection and measurement of sentiment at the concept level and provides high accuracy results.[57]

Conclusions

The new approach developed combines machine learning techniques (sentiment analysis and data mining) with multivariate analysis methods (SNA and text mining). Free software, that is very easy to access and use, has been used to do this. We are currently working on a research project aiming at integrating all these software tools into a Decision Support System, easier to use and interpret the results.

The sentiment analysis approach has proven its validity to evaluate the social mood of citizens in different time scales, registering the different clusters that emerged, gauging public states of mind via the collective valence and detecting the prevalence of the different emotions in the successive phases of the pandemic.

The approach has also shown, albeit rather indirectly, social support for public policies. Overcoming the conceptual limitations around the study of emotions may considerably enrich the perspectives and applications of sentiment analysis and similar kinds of studies, particularly thinking in the emerging mental pathologies—and not only in viral pandemics—around the 'information society'.

Finally, the combination in formal models of objective and subjective information, in this case about the COVID-19 vaccination process in Spain, will provide a more accurate vision of social reality, which will enable a more effective resolution of problems.

Author statements

Ethical approval

This work did not need to be approved by an ethics committee, as we used public information and messages from the social network Twitter.

Funding

This research is partially funded by the 'Grupo Decisión Multicriterio Zaragoza' research group (S35-20R) and the project 'Participación Ciudadana Cognitiva y Decisiones Públicas. Aplicaciones Sociosanitarias' (Ref. LMP35_21), both supported by grants from the Regional Government of Aragon and FEDER funds. The funders had no role in the design of the study; in the writing of the article, or in the decision to publish the results.

Competing interests

None declared.

Authors' contributions

A.T. contributed to conceptualisation, methodology, software, data curation, formal analysis, and writing, reviewing and editing the article. A.A. contributed to formal analysis and reviewing and editing the article. J.M.M.-J. contributed to conceptualisation, methodology, reviewing and editing, and funding acquisition. J.N. contributed to conceptualisation, methodology, software, data curation, formal analysis, and writing, reviewing and editing the article.

References

1. Sattar NS, Arifuzzaman S. COVID-19 vaccination awareness and aftermath: public sentiment analysis on twitter data and vaccinated population prediction in the USA. *Appl Sci* 2021 Jun 30;**11**(13):6128.
2. Yousefinaghani S, Dara R, Mubareka S, Papadopoulos A, Sharif S. An analysis of COVID-19 vaccine sentiments and opinions on Twitter. *Int J Infect Dis* 2021 Jul;**108**:256–62.
3. Kapoor KK, Tamilmani K, Rana NP, Patil P, Dwivedi YK, Nerur S. Advances in social media research: past, present and future. *Inf Syst Front* 2018 Jun 1;**20**(3):531–58.
4. Öztürk N, Ayvaz S. Sentiment analysis on Twitter: a text mining approach to the Syrian refugee crisis. *Telematics Inf* 2018 Apr;**35**(1):136–47.
5. https://twitter.com/; 2022.
6. https://www.statista.com/statistics/303681/twitter-users-worldwide/; 2022.
7. https://www.internetlivestats.com/twitter-statistics/; 2022.
8. Chandra Pandey A, Singh Rajpoot D, Saraswat M. Twitter sentiment analysis using hybrid cuckoo search method. *Inf Process Manag* 2017 Jul;**53**(4):764–79.
9. Zimbra D, Abbasi A, Zeng D, Chen H. The state-of-the-art in twitter sentiment analysis. *ACM Trans Manag Inf Syst* 2018 Sep. 5;**9**(2):1–29.
10. Naseem U, Razzak I, Musial K, Imran M. Transformer based deep intelligent contextual embedding for twitter sentiment analysis. *Future Generat Comput Syst* 2020 Dec;**113**:58–69.
11. Mendon S, Dutta P, Behl A, Lessmann S. A hybrid approach of machine learning and lexicons to sentiment analysis: enhanced insights from twitter data of natural disasters. *Inf Syst Front* 2021 Sep. 14;**23**(5):1145–68.
12. Giachanou A, Crestani F. Like it or not. *ACM Comput Surv* 2016 Nov 11;**49**(2):1–41.
13. Antonakaki D, Fragopoulou P, Ioannidis S. A survey of Twitter research: data model, graph structure, sentiment analysis and attacks. *Expert Syst Appl* 2021 Feb;**164**:114006.
14. Hasan MdR, Maliha M, Arifuzzaman M. Sentiment analysis with NLP on twitter data. In: *2019 international conference on computer, communication, chemical, materials and electronic engineering (IC4ME2)*. IEEE; 2019. p. 1–4.
15. Zahra K, Imran M, Ostermann FO. Automatic identification of eyewitness messages on twitter during disasters. *Inf Process Manag* 2020 Jan;**57**(1):102107.
16. Agarwal A, Singh R, Toshniwal D. Geospatial sentiment analysis using twitter data for UK-EU referendum. *J Inf Optim Sci* 2018 Jan 2;**39**(1):303–17.
17. Ilyas SHW, Soomro ZT, Anwar A, Shahzad H, Yaqub U. Analyzing Brexit's impact using sentiment analysis and topic modeling on Twitter discussion. In: *The 21st annual international conference on digital government research*. New York, NY, USA: ACM; 2020. p. 1–6.
18. Somula R, Dinesh Kumar K, Aravindharamanan S, Govinda K. *Twitter sentiment analysis based on US presidential election 2016*. 2020. p. 363–73.
19. Chaudhry HN, Javed Y, Kulsoom F, Mehmood Z, Khan ZI, Shoaib U, et al. Sentiment analysis of before and after elections: twitter data of U.S. Election 2020. *Electronics (Basel)* 2021 Aug 27;**10**(17):2082.
20. Budiharto W, Meiliana M. Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis. *J Big Data* 2018;**5**(1):51. C.
21. Sharma A, Ghose U. Sentimental analysis of twitter data with respect to general elections in India. *Proc Comput Sci* 2020;**173**:325–34.
22. Patel R, Passi K. Sentiment analysis on twitter data of world cup soccer tournament using machine learning. *IoT* 2020 Oct 10;**1**(2):218–39.
23. Asif M, Ishtiaq A, Ahmad H, Aljuaid H, Shah J. Sentiment analysis of extremism in social media from textual information. *Telematics Inf* 2020 May;**48**:101345.
24. Kasson E, Singh AK, Huang M, Wu D, Cavazos-Rehg P. Using a mixed methods approach to identify public perception of vaping risks and overall health outcomes on Twitter during the 2019 EVALI outbreak. *Int J Med Inform* 2021 Nov;**155**:104574.
25. Garcia K, Berton L. Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA. *Appl Soft Comput* 2021 Mar;**101**:107057.
26. Manguri K H, Ramadhan R N, Mohammed Amin R, Twitter P. Sentiment analysis on worldwide COVID-19 outbreaks. *Kurd J Appl Res* 2020 May 19:54–65.
27. Turón A, Aguaron J, Escobar MT, Moreno-Jiménez JM. Decision analysis in e-cognocracy using dynamic social networks. In: *ICDSST2020 conference*. Zaragoza; 2020.
28. Scott J. Social network analysis. *Sociology* 1988 Feb 2;**22**(1):109–27.
29. West R, Paskov HS, Leskovec J, Potts C. Exploiting social network structure for person-to-person sentiment analysis. *Trans Assoc Comput Linguist* 2014 Dec;**2**:297–310.
30. https://www.sanidad.gob.es/profesionales/saludPublica/prevPromocion/vacunaciones/covid19/docs/COVID-19_EstrategiaVacunacion.pdf; 2022.
31. https://www.sanidad.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov/vacunaCovid19.htm; 2022.
32. https://www.vacunacovid.gob.es/; 2022.
33. https://www.openstreetmap.org; 2022.
34. http://gephi.org; 2022.
35. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J Statist Mech Theory Exp [Internet]* 2008;**2008 Oct**(10):P10008. https://doi.org/10.1088/1742-5468/2008/10/p10008.
36. Jacomy M, Venturini T, Heymann S, Bastian M. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS One* 2014 Jun 10;**9**(6):e98679.
37. Jockers M. *Syuzhet: extracts sentiment and sentiment derived plot arcs from text, R package version 1.0-6*. https://cran.r-project.org/web/packages/syuzhet/syuzhet.pdf; 2020.
38. Bravo-Marquez F, Frank E, Mohammad SM, Pfahringer B. Determining word-emotion associations from tweets by multi-label classification. In: *2016 IEEE/WIC/ACM international conference on Web intelligence (WI)*. IEEE; 2016. p. 536–9.
39. Mohammad SM, Turney PD. Crowdsourcing a word-emotion association lexicon. *Comput Intell* 2013 Aug;**29**(3):436–65.
40. Mohammad SM. Sentiment analysis. In: *Emotion measurement*. Elsevier; 2021. p. 323–79.
41. de las Heras-Pedrosa C, Sánchez-Núñez P, Peláez JI. Sentiment analysis and emotion understanding during the COVID-19 pandemic in Spain and its impact on digital ecosystems. *Int J Environ Res Publ Health* 2020 Jul 31;**17**(15):5542.
42. Liang H, Fung ICH, Tse ZTH, Yin J, Chan CH, Pechta LE, et al. How did Ebola information spread on twitter: broadcasting or viral spreading? *BMC Publ Health* 2019 Dec 25;**19**(1):438.
43. Rufai SR, Bunce C. World leaders' usage of Twitter in response to the COVID-19 pandemic: a content analysis. *J Public Health* 2020 Aug 18;**42**(3):510–6.
44. Wang Y, Croucher SM, Pearson E. National leaders' usage of twitter in response to COVID-19: a sentiment analysis. *Front Commun (Lausanne)* 2021 Sep. 8:6.
45. Handel S. *Classification of emotions*. http://www.theemotionmachine.com/classification-of-emotions/; 2012.
46. SeshathriAathithyan S, Sriram Mv, Prasanna S, Venkatesan R. Affective — hierarchical classification of text — an approach using NLP toolkit. In: *2016 international conference on circuit, power and computing technologies (ICCPCT)*. IEEE; 2016. p. 1–6.
47. Russell JA, Barrett LF. Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *J Pers Soc Psychol* 1999;**76**(5):805–19.
48. Russell JA. A circumplex model of affect. *J Pers Soc Psychol* 1980;**39**(6):1161–78.
49. Barrett LF. Solving the emotion paradox: categorization and the experience of emotion. *Pers Soc Psychol Rev* 2006 Feb 21;**10**(1):20–46.
50. Ekman P. An argument for basic emotions. *Cognit Emot* 1992 May 7;**6**(3–4):169–200.
51. Ekman P. Basic emotions. In: Dalgleish T, Power M, editors. *Handbook of cognition and emotion*. Sussex (UK): John Wiley & Sons; 1999. p. 45–60.
52. Navarro J, Marijuán PC. The natural, artificial, and social domains of intelligence: a triune approach. *Proc West Mark Ed Assoc Conf* 2022. In press.
53. Bernal AI. Audiencias y temas de noticias en medios online y Twitter. El caso de elpais.com. In: *V congreso internacional de Ciberperiodismo y Web 20*. Audiencias Activas y Periodismo.; 2013.
54. Navarro J, Turón A, Altuzarra A, Lahoz-Beltra R. Comparative sentiment analysis of COVID-19: a machine learning approach. In: *7th ICDSST 2021 (international conference on decision support System technology)*. Loughborough (Reino Unido); 2021.
55. Verma B, Thakur RS. *Sentiment analysis using lexicon and machine learning-based approaches: a survey*. 2018. p. 441–7.
56. Catelli R, Pelosi S, Esposito M. Lexicon-based vs. Bert-based sentiment analysis: a comparative study in Italian. *Electronics (Basel)* 2022 Jan 26;**11**(3):374.
57. D'Andrea A, Ferri F, Grifoni P, Guzzo T. Approaches, tools and applications for sentiment analysis implementation. *Int J Comput Appl* 2015;**125**(3).

Public Health 215 (2023) 17–26

Contents lists available at ScienceDirect

Public Health

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

Review Paper

Impact of economic growth on physical activity and sedentary behaviors: a Systematic Review

H. Yang [a], R. An [b], C.V. Clarke [c], J. Shen [d, *]

[a] Department of Physical Education, Capital University of Economics and Business, China
[b] Brown School, Washington University in St. Louis, United States
[c] Department of Kinesiology and Community Health, University of Illinois Urbana-Champaign, United states
[d] Department of Physical Education, China University of Geosciences (Beijing), China

ARTICLE INFO

Article history:
Received 29 June 2022
Received in revised form
20 October 2022
Accepted 25 November 2022
Available online 4 January 2023

Keywords:
Macroeconomic growth
Economic development
Physical activity
Sedentary behavior
Review

ABSTRACT

Objectives: This study systematically reviewed scientific evidence regarding the impact of economic growth on physical activity (PA) and sedentary behaviors.
Methods: A keyword and reference search was conducted in PubMed, Web of Science, Cochrane Library, Scopus, and EBSCO from the inception of an electronic bibliographic database to November 2021. Studies that met all of the following criteria were included in the review: (1) study designs: observational studies; (2) study subjects: people of all ages; (3) exposure: macroeconomic growth; (4) outcomes: PA or sedentary behaviors/sitting time.
Results: 15 studies were identified. Eight among 12 studies found economic growth positively associated with at least one domain/measure of PA, whereas the remaining four found an inverse relationship. One of two studies examined the association of economic growth and physical inactivity found a positive relationship, while another found no significant relationship. Four studies examined the associations between economic growth and sedentary behaviors, and the results were inconclusive. The impact of economic growth on PA is through three main pathways—(a) building and maintaining parks and green spaces, (b) adoption of the modern workplace and high technology, and (c) motorized transportation.
Conclusion: Building and maintaining parks and green spaces can effectively promote leisure time physical activity and active commute in developed countries, whereas promoting workplace and transportation-related PA could be prioritized in developing nations. Future research calls for longitudinal study design and further exploration of macro-environmental factors. Policymakers and stakeholders should be informed about the potential PA reduction resulting from economic growth and develop preventive strategies to alleviate the problem.

Introduction

Physical activity (PA) is essential for the prevention and treatment of non-communicable diseases, such as cardiovascular diseases, cancers, and diabetes.[1] Additionally, PA has positive effects on mental health, the delayed onset of dementia, and a healthy weight status.[1–6] However, a growing economy, urbanization, and motorized transport have decreased PA[7,8,.9] Sedentary behavior is a leading risk factor for chronic morbidity and global mortality.[10] Approximately, a quarter of all adults and 81% of adolescents globally did not achieve sufficient PA in 2016, with over 1.4 billion adults at a risk of developing or exacerbating diseases linked to inactivity.[1,11] The prevalence of insufficient PA in adolescents varied from 18.7% to 90.6%, with a median of 79.7%.[12] Recognizing the importance and urgency of reducing global levels of insufficient PA, member states of the World Health Organization agreed on a global target of a 15% increasing PA by 2030.[13] In addition to insufficient engagement in PA, many adults spend much of their waking time sitting.[14] Sedentary behavior refers to certain activities characterized by an energy expenditure $\leq$1.5 metabolic equivalents, while in a sitting, reclining, or lying posture.[15] Sedentary behavior increased by 5% (from 31.6% to 36.8%) in high-income countries between 2001 and 2016.[16] Sedentary behaviors are related to an increased risk for chronic diseases and premature mortality.[17] Furthermore, even individuals meeting the recommended weekly PA may still

* Corresponding author.
E-mail address: shenjing@cugb.edu.cn (J. Shen).

https://doi.org/10.1016/j.puhe.2022.11.020
0033-3506/

have high levels of sedentary behaviors, which increases the health risks.[18]

To promote PA and reduce sedentary time, insights into the factors influencing these behaviors are needed. Various factors across the social-ecological levels have been identified[19–21], including workplaces, neighborhoods, regions, and countries. The vast majority of previous studies have considered the relationship between neighborhood characteristics and PA and sedentary behaviors.[22–26] More recently, research has shifted toward macro-environment factors,[27,28] including how macro-environmental changes such as economic growth have influenced PA and sedentary behaviors.[8,29] Economic growth, one of two global megatrends,[30] has dramatically changed the environment in which modern humans live and work. The impact of this change in the living environment on PA-related behaviors and, consequently, health outcomes (overweight, obesity, and chronic diseases) has attracted much interest in recent years. Economic growth will influence the availability of natural and human-created resources. First, it may facilitate the provision of sport facilities, parks, and green space. Second, it is linked to the modern workplace and high technology, which may increase the occupation transition from labor-intensive jobs to sedentary jobs. Third, it may increase the motorized transportation and accessibility to technological household and entertainment appliances, thereby decreasing PA and stimulating sedentary behaviors. Dancause et al.[31] reported that the economic development in Vanuatu is accompanied by increased sedentary recreation, although PA levels remain high. Dumith et al.[9] reported economic progress does not necessarily increase PA levels. In Europe, Cameron et al.[27] found that higher macroeconomic indicators were associated with leisure time physical activity (LTPA) but not with total PA. However, that review used the data only from previous pan-European PA studies, with no evidence available regarding the rest of the countries worldwide.

This current study aimed to review and synthesize the empirical evidence on the association between economic growth and PA. One essential contribution of this review is the hypothesized pathways linking the economic growth to PA worldwide and thereby provides a template for the future RCTs or cohort studies. Findings from this review can be informative to policymakers and stakeholders concerned with policy assessment, help develop strategies that counteract economic growth-related decreases in PA or increase in sedentary behaviors, and effectively alleviate the adverse health outcomes such as non-communicable diseases and cardiovascular diseases.

Methods

This review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.[32]

Study selection criteria

Studies that met all of the following criteria were included in the review: (a) study designs: observational studies, (b) geographical coverages: any country or region worldwide, (c) exposure: economic growth (e.g. gross domestic product (GDP), GDP per capita, purchasing power parity per capita or PPP), (d) outcomes: prevalence of PA/physical inactivity, sitting time, sedentary behavior, (e) article type: peer-reviewed publications, (f) time window of search: from the inception of an electronic bibliographic database to November 30, 2021, and (g) language: articles written in English.

Search strategy

A keyword search was performed in five electronic bibliographic databases: PubMed, Web of Science, Cochrane Library, Scopus, and EBSCO, and a hand search of cross references. The complete list of all possible combinations of keywords is provided in Appendix A. The complete list of keywords and search algorithms in PubMed is provided in Appendix B.

Data extraction and synthesis

A standardized data extraction form was used to collect the methodological and outcome variables from each included study (Tables 1–3). Heterogeneities in the study design and measures of economic growth (e.g. GDP, GDP per capita, PPP, and real GDP) precluded a meta-analysis. We thus summarized the common themes and findings of the included studies narratively. Two co-authors of this review independently conducted the data extraction, theme identification, and narrative summarization. Discrepancies were resolved through discussion under the participation of a third co-author.

Study quality assessment

We used the National Institutes of Health's Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies to assess the quality of each included study. The study quality assessment helped to measure the strength of scientific evidence but was not used to determine the inclusion of studies. Two co-authors of this review independently conducted the study quality assessment, with discrepancies resolved through discussion with a third co-author.

Results

Description of the studies

A total of 15 articles met the eligibility criteria and were included in the review[33–47](Fig. 1). Nine of the 15 studies assessed the relationship between economic growth and PA or physical inactivity at the country-level (i.e. using country-level aggregate statistics as the unit of analysis), whereas the remaining six analyzed city/region/province-level data[36,41,43–45,47](Table 1). Table 2 summarizes the data sources and measures for economic growth and PA among the studies included in the review. Data on economic growth were mostly obtained from the World Development Indicators of the World Bank (n = 6). Included studies mainly focused on GDP (n = 4), GDP per capita (n = 7) in relation to the PA.

Effects and main findings of economic growth on PA

Table 3 reports the key findings on the relationship between economic growth and PA. 12 studies examined the association between economic growth and PA. Among them, eight studies found economic growth to be positively associated with at least one domain or measure of PA, while the remaining four found inverse associations. One of two studies examined the association of economic growth and physical inactivity found a positive relationship, while another found no significant relationship. Four studies examined the association of economic growth and sedentary behavior, and the results were inconclusive. Some but not all studies explored specific mechanisms linking economic growth to PA, which were summarized into three major pathways.

Table 1
Basic characteristics of the studies included in the review.

Study ID	Author, Year	Unit of analysis	Countries/regions studied	Sample size	Study design	Age	Gender	Statistical model	Study period	Country characteristics
1	Bosdriesz,2012	Country	38	177,035	Cross-sectional	18–69	Both	Multilevel logistic regression	2002–2005	LMICs
2	Van Tuyckom,2011	Country	27	24,846	Cross-sectional	15+	Both	Bivariate linear regression models	2005	EU Countries
3	Haase,2004	Country	23	19,298	Cross-sectional	17–30	Both	Multilevel logistic regression	1999–2001	HICs and LMICs
4	Wang,2017	Region	30	43,389	Cross-sectional	20–69	Both	Logistic regression	2013–2014	UMICs
5	Atkinson, 2016	Country	47	196,742	Cross-sectional	20–69	Both	Multilevel logistic regression	2002–2003	LMICs
6	Ruseski, 2014	Country	34	49,730	Cross-sectional	15–98	Both	Probit model	2006–2008	HICs and LMICs
7	Ma, 2020	Country	68	180,298	Cross-sectional	12–15	Both	Linear regression; Logistic regression	2009–2016	LMICs
8	Linando, 2018	Country	3		Longitudinal		Both	Path analysis	1980–2015	HICs and UMICs
9	Van Cauwenberg,2018	Region	24	51,820	Repeated Cross-sectional	18+	Both	Multilevel logistic regression	2002,2005,2013	EU Countries
10	Lee, 2003	Country	1		Longitudinal		Both	Path analysis	1969–1993	HIC
11	Chen,2016	City	5	13,950	Cross-sectional	13.6 ± 1.0	Both	Multilevel hierarchical analysis	2009–2010	UMICs
12	Pascual, 2007	Province	1	19,324	Cross-sectional	16–74	Both	Multilevel logistic regression	1980,1990,2000	HIC
13	Wicker, 2017	Region, individual	28	18,675	Cross-sectional	18–64	Both	Multilevel logistic regression	2013–2015	HICs and LMICs
14	Werneck,2020	Country	7	128,602	Cross-sectional	18–64	Both	Multilevel logistic regression	2010–2016	LMICs
15	Yang,2017	Individual	1	9487	Longitudinal	6–17	Both	Logistic regression	1997–2011	UMIC

Notes: LMICs, Low and Middle-Income Countries; EU, European Union; HICs, High-Income Countries; UMICs, Upper Middle-Income Countries.

Relationship between economic growth and PA

Van Tuyckom et al.[34] reported significantly more LTPA in European countries with a higher GDP and real GDP. Haase et al.[35] documented that the likelihood of LTPA was positively associated with national economic development (per capita GDP). Wang et al.[36] found that people in economically advanced regions in China currently engage in more physical activities than those in less economically developed regions. Ruseski et al.[38] reported that higher GDP per capita is associated with higher PA participation. Werneck et al.[46] reported that the higher GDP was associated with higher LTPA, and higher transport PA, not associated with total PA among South American adults. Wicker et al.[45] reported that regional GDP per capita was positively related to individual PA level meeting and exceeding the World Health Organization (2010) guidelines in 28 European countries. Ma et al.[39] reported a weak positive association between a country's PPP/capita and the level of PA among adolescents. For different subgroups, Van Cauwenberg et al.[41] reported women, those with at least 20 years of education, white-collar workers or self-employed, retirees, and students were more likely to achieve PA guidelines in areas with increased GDP.

In contrast, Linando et al.[40] reported PA transition was significantly associated with socio-economic transitions. The nations included in the study showed a significant occupational shift from agrarian to service sector with the increase of GDP. Subsequently, the amount of occupational PA decreased. Lee et al.[42] found that South Koreans tended to decrease their PA with the socio-economic transition. Yang et al.[47] reported that children from households with lower income were more likely to engage in active commuting to school (ATS). Bosdriesz et al.[33] reported that GDP was negatively correlated with the population's PA levels.

Two studies examined the association of economic growth and physical inactivity. Atkinson et al.[37] found no statistically significant association between physical inactivity and economic development, but individuals' income was positively associated with physical inactivity. Pascual et al.[44] found that the provinces with the lowest GDPpc were associated with fewer sports facilities and had the highest prevalence of physical inactivity.

Relationship between economic growth and sedentary behavior

The effects of economic development on sedentary behavior/sitting time remained mixed. Chen et al.[43] documented that higher levels of economic growth (as indicated with GDP per capita) were associated with longer sitting hours. Van Cauwenberg et al.[41] reported that GDP was significantly positively related to the odds of high levels of sitting in the total sample. By contrast, Ma et al.[39] reported an inverse association between a country's PPP and sedentary behavior. Van Cauwenberg et al.[41] reported that an increase in GDP was significantly negatively related to sitting time among students. Werneck et al.[46] documented that GDP was not associated with sitting time.

Study quality assessment

Table 4 reports criterion-specific and global ratings from the study quality assessment. The studies included in the review on average scored 6.3 out of 14 (range: 5–8). All included studies clearly stated the research question/objective, specified and defined the study population, had a participation rate of ≥50%, recruited participants from the same or similar populations during the same time period, pre-specified and uniformly applied inclusion and exclusion criteria to all potential participants, and implemented the valid and reliable exposure measures. Most studies measured and statistically adjusted key potential confounding variables for their impact on the relationship between exposures and outcomes (n = 9). In contrast, none of the studies had the

Table 2
Data sources and measures of economic growth and physical activity in the studies included in the review.

Study ID	Data source of economic growth	Detailed measure of economic growth	Data source of physical activity	Detailed measure of physical activity
1	World Bank for 2002, with the exception of Myanmar (2007)	GDP	WHS(2011)	Vigorous PA, moderate PA and walking
2	HFA	GDP; Real GDP	Eurobarometer survey	Leisure Time PA (LTPA)
3	World Bank development indicators	GDP per capita	IHBS	The odds of any leisure PA, the odds of recommended frequency PA
4	National Bureau of Statistics of China and provincial level statistics bureaus (2013)	GDP per capita	2014 physical activity and physical fitness survey in China	Weekly physical activity
5	WHS (2002–2003)	GDP per capita; Individual income	WHS 2003	1.The frequency and duration of PA: vigorous-, moderate-, and low-intensity PA levels 2.Physical inactivity
6	World Bank Online Database for 2007	GDP per capita (in U.S.2009 PPP dollars); Individual income	ISSP 2007	The probability of PA participation
7	World Bank (60 courtries; Index Mundi (8 countries)	PPP/capita	GSHS	Moderate to Vigorous intensity physical activity
8	World Bank databank	GDP	World Bank databank	PA Transition
9	Eurostat database	GDP	Special Eurobarometer survey	1.Vigorous PA, moderate PA and walking 2. Sitting time
10	Korean Statistical Yearbook	GNP	Korean Statistical Yearbook; Statistical yearbook of Transportation	PA transition: Industrial activity; Transportation
11	World Bank 2009 database; IMF 2009; National bureau of Statistics of China 2009	GDP per capita	CSHS	Sitting time, sedentary activities
12	Eurostat database (1980, 1990, 2000)	Provincial GDP per capita	National Health Survey (2001)	Physical inactivity
13	Eurostat (2013)	Regional GDP per capita	Eurobarometer (2013)	Sport and PA levels
14	1.Argentina: Instituto Nacional de Estadística y Censos 2.Brazil: Instituto Brasileiro de Geografia e Estatística 3.Chile: Instituto Nacional de Estadísticas 4.Colombia: Departamento Administrativo Nacional de Estatística 5.Ecuador: Instituto Nacional de Estadística y Censos and Pontificia Universidad Católica del Ecuador 6.Peru: Instituto Nacional de Estadística e Informática 7.Uruguay: Instituto Nacional de Estadística	GDP	Self-reported questionaires: IPAQ, GPAQ	Leisure Time PA, Transport PA,Occupational PA, total PA, sitting
15	CHNS	Per capita household income	CHNS	ATS

Notes: GDP, Gross Domestic Product; GNP, Gross National Product; ATS, Active Travel To School; CHNS, The China Health and Nutrition Survey; CSHS, The Chinese Student Health Survey; HFA, European Health for All Database, World Health Organization; WHS, World Health Survey; PA, physical activity; IHBS, International of Health and Behaviour Survey; ISSP, International Social Survey Programme; GSHS, Global School-based Student Health Survey; IPAQ, The International Physical Activity Questionnaire; GPAQ, the Global Physical Activity Questionnaire.

Table 3
Estimated effects and main findings of economic growth on physical activity in the studies included in the review.

Study ID	Estimated effect of economic growth on physical activity	Main findings
1	1. GDP was negative associated with men's moderate PA (OR = 0.76, 95% CI: 0.65, 0.89), and walking (OR = 0.79, 95% CI: 0.63, 0.98). 2. GDP was associated with moderate PA for the white-collar group (OR = 0.81, 95% CI: 0.69, 0.96), those not working for pay (OR = 0.82, 95% CI: 0.67, 1.00), the blue collar group (OR = 0.87, 95% CI: 0.72, 1.04), and agricultural workers (OR = 0.91, 95%CI: 0.76, 1.09).	1. A negative association was found between gross domestic product and PA. 2. Economic development is one of the determinants of the level of overall PA at national-levels.
2	1. GDP was associated with overall, male, and female Leisure Time Physical Activity (LTPA), respectively (β = 0.599, 90%CI: 2.272, 6.083; β = 0.65, 90%CI: 2.482, 5.781; β = 0.552, 90%CI: 2.009, 6.293). 2. Real GDP was significantly associated with overall, male, and female LTPA, respectively (β = 0.626, 90%CI: 4.732, 11.731; β = 0.682, 90%CI: 5.170, 11.162; β = 0.577, 90%CI: 4.225, 12.136).	There were significant associations between overall LTPA and GDP, real GDP.
3	1. The prevalence of LTPA at any level was positively correlated with economic development (r = 0.49, $P = 0.02$). 2.The likelihood of being physically active and physical activity at recommend level were greater in respondents from more economically developed nations, respectively (OR 1.38, 95%CI 1.33−1.43; OR 1.21,95%CI 1.15−1.27).	The likelihood of LTPA was positively associated with national economic development (per capita gross domestic product).
4	1.The correlation between GDP per capita and weekly activity levels in males was r = 0.23 and r = 0.15 in females. 2.Participants from higher economically advanced region have increased physical activity levels than their counterparts in less economically advanced regions ($P < 0.01$). 3.Compared with those in less economically advanced regions, the participants in higher economically advanced regions have higher tendencies to not reach adequate activity levels ($P < 0.01$).	People in economically advanced regions in China currently engage in more physical activities than those in less economically developed regions.
5	1.Individuals living in urban areas were 27% more likely to be physically inactive than individuals living in rural areas (OR: 1.27, 95%CI: 1.23−1.32). Individuals with an income of Quintile one were 17% less likely to be physically inactive compared with individuals with an income of Quintile 5 (OR: 0.83, 95%CI: 0.79 −0.87). 2.The association between each country-level variable was assessed separately, while adjusting for individual factors. Neither economic development nor urbanization was statistically significant.	1.Individual income was positively associated with physical inactivity. As countries experience economic development, changes are also seen in their occupational structure shifts from agrarian to industrial-based, which result in increased countrywide physical inactivity levels simultaneously. 2.There was no statistically significance between variable economic development and physical inactivity.
6	1.There was a positive relationship between income and the probability that an individual participates in physical activity. 2.The higher GDP per capita in a country, the more likely are individuals to report participating in sport and physical activity.	Gross domestic product per capita is associated with higher physical activity participation.
7	1.PPP/capita was positively associated with the number of days with sufficient physical activity and duration of sedentary behavior per day, respectively. 2.Compared to the second quintile of PPP, the odds of having sufficient physical activity were larger in the upper PPP/capita categories (Q3: OR = 1.31, 95% CI = 1.15−1.50; Q4: OR = 1.39, 95% CI = 1.20−1.61; Q5: OR = 1.52, 95% CI = 1.34 −1.73). 3.compared to the first PPP/capita quintile, the odds of having low sedentary time were also higher in the upper PPP/capita categories (Q2: OR = 0.44, 95% CI = 0.38−0.51; Q3: OR = 0.51, 95% CI = 0.44−0.60; Q4: OR = 0.18, 95% CI = 0.16 −0.21; Q5: OR = 0.23, 95% CI = 0.20−0.26).	There was a weakly positive association of a country's PPP/ capita with the level of PA among adolescents and the inverse association between a country's PPP and sedentary behavior.
8	GDP had a strong influence on physical activity transition (−0.88 ~ -0.93).	Physical activity transition was significantly associated with socio-economic transitions. All three nations included in the study showed significant decrease in percentage of the population employed in agriculture and significant increase in percentage of the population in service sector.
9	1.In the total sample, GDP was not significantly related to the odds of meeting PA guidelines cross-sectionally. 2.In the total sample, a significant positive relationship was observed longitudinally; an increase in GDP of 10,000 euro/inhabitant within a region was related to a 1.49 times higher odds of meeting PA guidelines. 3.An increase in GDP of 10,000 euro/inhabitant was related to a significantly stronger increase in the odds of meeting PA guidelines among women (OR = 1.60, 95% CI = 1.19−2.16) compared to men (OR = 1.36, 95% CI = 1.00 −1.85). 4.An increase in GDP of 10,000 euro/inhabitant was related to a 1.96, 1.68, 1.54 and 2.41 times higher odds of meeting PA guidelines among those with at least 20 years of education, white-collar workers or self-employed, retirees and students, respectively. 5.Cross-sectionally, GDP was significantly positively related to the odds of high levels of sitting in the total sample and in all subgroups, except for women, those aged 65 years and older, manual workers and retirees. 6.Longitudinally, GDP was not significantly related to sitting time in the total sample, but it was significantly negatively related to the odds of high levels of sitting time among students.	1.An increase in GDP was more strongly related to a higher odds of meeting PA guidelines among women than men. Furthermore an increase in GDP was related to a higher odds of meeting PA guidelines among those with at least 20 years of education, white-collar workers or self-employed, retirees and students, but not among other education or occupation subgroups. 2.Cross-sectionally, GDP was significantly positively related to the odds of high levels of sitting in the total sample; Longitudinally, an increase in GDP was significantly negatively related to sitting time among students, but no relationships were observed among other subgroups.
10		South Koreans tended to decrease their physical activity in relationship to the socio-economic transition.

(*continued on next page*)

Table 3 (*continued*)

Study ID	Estimated effect of economic growth on physical activity	Main findings
	1.In South Korea, the GNP increased more than 50-fold over the 24 years, while physical activity decreased in job-related and transportation areas. 2.GNP had a strong influence on the proportion of tertiary industry.	
11	GDP was positively associated with hours of sitting (Coefficient = 0.4278, $P < 0.01$) .	Higher levels of economic development (as indicated with GDP per capita) were associated with longer sitting hours.
12	The odds ratio in residents of provinces with the lowest current GDPpc versus those with the highest was 1.64 [95% confidence interval (CI): 1.17–2.30] in men and 2.01 (95% CI: 1.48–2.73) in women. The odds ratio in residents of provinces that had always been among those with the lowest GDPpc versus residents of provinces that had never been among those with lowest GDPpc was 1.54 (95% CI: 1.18–2.00) in men and 1.91 (95% CI: 1.51–2.41) in women.	The provinces with the lowest GDPpc and those that have repeatedly had the lowest GDPpc over the two preceding decades are associated with fewer sports facilities and have the highest prevalence of physical inactivity.
13	Regional GDP per capita was positively associated with individual PA level meeting (Pass) (OR = 1.00, $P < 0.01$) and exceeding the guidelines (Extra) (OR = 1.00, $P < 0.001$).	Regional government quality has a significant and positive association with individual participation in sport and physical activity at a level meeting or exceeding the guidelines. The impact is much larger than that of regional GDP per capita, indicating that regional disadvantage in terms of political quality is more relevant than being disadvantaged in terms of economic wealth.
14	1.Higher GDP was associated with higher transport PA [ORwomen: 1.39 (95% CI: 1.20–1.61); ORmen: 1.16 (95% CI: 1.00–1.35)], with low variation (I2: 0% in both sexes). 2.The association between the highest tertile of gross domestic product and leisure physical activity overlapped the unit in both sexes [ORmen: 1.16 (95% CI: 0.94–1.42); ORwomen:1.12 (95% CI: 0.97–1.29)], with a substantial heterogeneity among men (I2: 65.2%) and a trivial variation among women (I2: 16.0%).	1.The higher GDP was associated with higher transport PA and leisure PA. 2.GDP was not associated with total PA. 3.No consistent associations were found for occupational physical activity and sitting time.
15	Children from households with low-income level [(OR:1.31 (95% CI:1.04, 1.65)], low parental education status [(OR:1.36 (95% CI:1.11, 1.65)], or no private vehicle [OR:1.54 (95% CI:1.29, 1.84)] were more likely to engage in ATS. There were no apparent differences in ATS between boys and girls.	Children from households with lower income level were more likely to engage in ATS.

outcome assessors blinded to the exposure status of the participants. None of the studies provided a sample size justification using power analysis. None of the studies measured the exposures of interest before the outcomes were measured. None of the studies implemented valid and reliable outcome measures. A total of four studies had a reasonably timeframe that was sufficient for changes in outcomes to be observed, four assessed the exposures more than once during the study period, two studies examined different levels of the exposure in relation to the outcome. Only one study had an attrition rate of ≤20%.

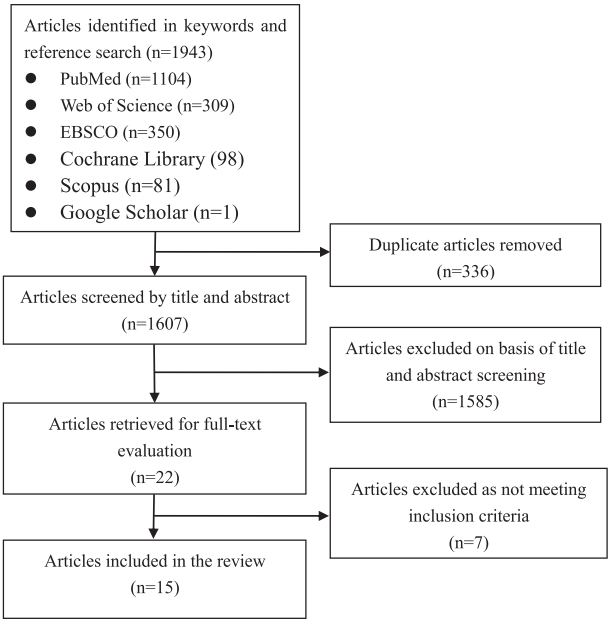


Fig. 1. Study selection flowchart.

Hypothesized pathway

Fig. 2 illustrates the hypothesized pathway of economic growth and PA. The levels of economic development influence the different domains of PA through three main pathways—(a) building and maintaining parks and green spaces, (b) adoption of modern workplace and high technology, and (c) motorized transportation. (a) Economic growth is associated with greater provision and maintenance of parks and green spaces, which are more conducive to LTPA. (b) Economic growth facilitates the adoption of the modern workplace and high-technologies, which resulted in occupational changes from labor-intensive to sedentary service-oriented professions.[37] (c) Economic growth is accompanied by motorized transportation, which decreases the popularity of walking and riding bicycle and cultivates the preference and demand for driving. Importantly, it should be noted that the pathways were hypothesized in the literature (e.g. in the introduction and/or discussion sections of an included article) but not empirically examined by any data-driven analytic approach.

Discussion

This study reviewed the scientific evidence regarding the associations between economic growth and PA. 15 studies that met the selection criteria were included in the review. PA behaviors included PA, LTPA, occupational PA, transport PA, physical inactivity, sedentary behavior/sitting time, and ATS. Twelve studies explored the associations between economic growth and PA. Among them, eight studies empirically confirmed a positive correlation between economic growth and the different domains of PA, especially LTPA,[34–36,38,39,41,45,46] whereas four studies reported a negative correlation between GDP and PA levels.[33,40,42,47] Two studies examined the association of economic growth and physical inactivity. One study reported a positive correlation between GDP and the prevalence of physical inactivity,[44] while another reported no significant association between physical inactivity and economic

Table 4
Study quality assessment.

Criterion	Study ID														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Was the research question or objective in this paper clearly stated?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
2. Was the study population clearly specified and defined?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3. Was the participation rate of eligible persons at least 50%?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
4. Were all the subjects selected or recruited from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study pre-specified and applied uniformly to all participants?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
5. Was a sample size justification, power description, or variance and effect estimates provided?	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	N	N	N	N	N	N	N	Y	Y	Y	N	N	N	N	Y
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?	N	N	N	N	N	N	Y	N	N	N	N	Y	N	N	N
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
10. Was the exposure(s) assessed more than once over time?	N	N	N	N	N	N	N	Y	Y	Y	N	N	N	N	Y
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
12. Were the outcome assessors blinded to the exposure status of participants?	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
13. Was loss to follow-up after baseline 20% or less?	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Y	N	Y	N	Y	Y	Y	Nik	N	N	Y	Y	Y	Y	N
Total score	6	5	6	5	6	6	7	7	7	7	6	7	6	6	8

***Notes*:** This study quality assessment tool was adopted from the National Institutes of Health's Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies. For each criterion, a score of one was assigned if "Yes" was the response, whereas a score of zero was assigned otherwise. A study-specific global score, ranging from zero to 14, was calculated by summing up scores across all 14 criteria. Study quality assessment helped measure strength of scientific evidence, but was not used to determine the inclusion of studies.

development.[37] Four studies explored the association of economic growth and sedentary behaviors, but the findings were inconsistent. Two studies reported a negative correlation between GDP and the prevalence of sedentary behaviors,[39,41] while two reported a positive association.[41,43] The mixed findings could be attributed to the differences in PA domains or measures assessed and variations in the economic development status of the countries studied.

The impact of economic growth on the different domains of PA and sedentary behaviors through three main pathways—(a) building and maintaining parks and green spaces, (b) adoption of modern workplace and high technology, and (c) motorized transportation. The positive association between economic growth and participation in PA could reflect either supply side- or demand side-effects in economically advanced countries. The positive association reflects supply side-effects to the extent that countries with higher GDP can build and maintain parks and green spaces,[48] which may contribute to higher levels of PA,[49–51] especially LTPA. The positive association reflects demand side-effects to the extent that countries with higher GDP enable individuals to have more control over their leisure time, thus increasing demands for fitness among

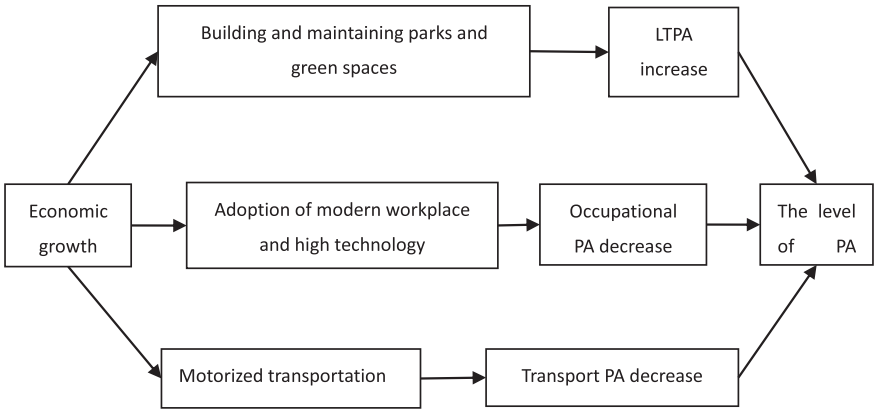


Fig. 2. The hypothesized pathway of economic growth and PA. PA, physical activity.

residents. Furthermore, individuals in countries with higher GDP per capita had higher disposable income to pay for sports and exercise participation. Thus, individuals are more likely to engage in LTPA in developed countries. However, individuals from low socioeconomic backgrounds may not benefit much due to a lack of leisure time or disposable income. Therefore, policymakers may need to consider that the increased prosperity benefits the PA levels of all residents.

As countries experience economic development, occupational structure shifts from agricultural to industrial occupations. This change involves higher mechanization, technology, and urbanization, bringing mechanized labor, sedentary occupations, and motorized transport.[37] In contrast, more activity is undertaken at work and for transportation for low-income or middle-income countries. Thus, the amount of occupational and transport PA in the population decreased.[37,52] Simultaneously, the population still lacks the financial and temporal resources to engage in LTPA.[37] Even individuals in white-collar occupations, who typically have more financial resources and leisure time to participate in LTPA, will experience an offset of PA due to a lower amount of occupational and transport-related PA, leading to an overall decline of countrywide PA. In addition, the inconsistent association between economic growth and sitting time[41,46] appears to be related to individual-level determinants differences.[10]

For children and adolescents, there was a weakly positive association between economic growth and participation in PA among children and adolescents and the inverse association between economic growth and sedentary behaviors.[39] The positive association reflects supply side-effects to the extent that countries with higher GDP can provide the facilities for LTPA, which may be more limited in schools and communities from less affluent countries. On the other hand, the inverse association reflects the accessibility to the technological household and entertainment appliances with higher GDP countries, which may be associated with higher level of screen-based sedentary behavior or sedentary socializing activities.[39] Over all, it is imperative to develop specific measures to promote PA and reduce sedentary behaviors among adolescents. It includes ATS, minimum number of hours PA in school, sports in communities, more open green spaces.

It is worth noting that physical inactivity and sedentary behaviors have significant adverse effects on public health.[6,17,53] Therefore, preventive strategies aimed at PA promotion and SB reduction should be developed. Certain regulations and policies should be in place to mitigate the possible physical inactivity and sedentary behaviors resulted from economic growth. The national-level policy is needed to encourage PA modes of transportation and promote participation in active recreation and sports in leisure time worldwide. Effective policies include the improved provision of cycling and walking infrastructure, improving road safety, and creating more opportunities for PA in public open spaces and parks, in workplaces, and in other local community settings.[6,54] Meanwhile, it is necessary to design the intervention measures which contribute to decrease the frequency and duration of sedentary throughout the day substituting with MVPA.[55] In sum, policy makers should closely monitor the prevalence of physical inactivity and sedentary behaviors in the population during the era of economic development and consider designing and implementing customized policies to help people maintain PA engagement and reduce sedentary behaviors.

To the best of our knowledge, this review serves as the first attempt to synthesize scientific literature regarding the impact of economic growth on PA and sedentary behaviors. Its strengths included large-scale population-based studies and the availability of data on PA across many countries. However, several limitations pertaining to this review and the selected studies should be considered. No studies included in this review reported quantitative estimates focusing on the same economic development and PA measures, which precluded meta-analysis. Most studies relied upon the cross-sectional design, and the heterogeneity of the PA measures prevented a causal interpretation regarding the impact of economic growth on PA. Observational studies-based association were prone to confounding bias. It could not be ruled out for the reverse causality in the hypothesized pathways when it might instead be that the PA level influences economies.[56] Therefore, the present review only constructed the hypothesized pathways. In addition, country-level studies using aggregated data were prone to ecological fallacy.[57]

Regarding measurements, the studies included in the review all relied on self-reported data in the analysis despite their limitations.[6,58] The literature search identified articles written in English only and excluded articles written in other languages. The scope of this review is limited to the literature that directly assessed the relationship between economic growth and PA. In contrast, the much broader literature concerning the impact of the economy (urbanization, human development, economic inequality, etc.) on PA was not reviewed.

Conclusion

This study reviewed and synthesized the scientific literature on the relationship between economic growth and PA. A total of 15 studies met the predetermined eligibility criteria and were included in the review. Twelve studies examined the association

between economic growth and PA. Among them, eight studies included in the review found economic growth to be positively associated with at least one domain and/or measure of PA, whereas the remaining four found inverse associations. One of two studies examined the association of economic growth and physical inactivity found a positive relationship, while another found no significant relationship. Four studies examined the association of economic growth and sedentary behaviors, and the results were inconclusive. Longitudinal study designs are warranted to identify the potential pathways linking economic growth to PA in the future. The findings of the present review may be informative to policymakers and stakeholders to formulate and implement policies which could help to alleviate the risk of people's decreased PA engagement and prolonged sedentary time resulting from economic growth.

Author statements

Acknowledgment

This study was financially supported by the Scientific Research Promotion Plan of ** University.

Ethical approval

The study did not involve human participants; therefore, it did not require ethics committee approval.

Funding

None declared.

Competing interests

All authors of this article declare that they have no conflicts of interest.

References

1. World Health Organization. Health topics: physical activity. https://www.who.int/health-topics/physical-activity#tab=tab_1. Accessed Feb 12, 2022.
2. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U. Global physical activity levels: surveillance progress, pitfalls, and prospects. *Lancet (London, England)* 2012;**380**(9838):247–57.
3. Sallis JF, Bull F, Guthold R, Heath GW, Inoue S, Kelly P. Progress in physical activity over the Olympic quadrennium. *Lancet (London, England)* 2016;**388**(10051):1325–36.
4. Hlth IISPA. The bangkok declaration on physical activity for global health and sustainable development. *Br J Sports Med* 2017;**51**(19):1389–91.
5. World Health Organization. Be he@lthy, be mobile: a handbook on how to implement mobile health for physical activity. https://www.who.int/publications/i/item/9789240033474. Accessed January 22, 2022.
6. Who. *Global action plan on physical activity 2018-2030: more active people for a healthier world*. World Health Organization; 2018.
7. Alessio H, Bassett D, Bopp M, Parr B, Patch G, Rankin J. Climate change, air pollution, and physical inactivity: is active transportation part of the solution? *Med Sci Sports Exerc* 2021;**53**(6):1170–8.
8. Van Holle V, Deforche B, Cauwenberg J, Goubert Liesbet, Maes L, Van de Weghe N. Relationship between the physical environment and different domains of physical activity in European adults: a systematic review. *BMC Publ Health* 2012;**12**:807.
9. Dumith SC, Hallal PC, Reis RS, Kohl HW. Worldwide prevalence of physical inactivity and its association with human development index in 76 countries. *Prev Med* 2011;**53**(1–2):24–8.
10. O'Donoghue G, Perchoux C, Mensah K, Lakerveld J, van der Ploeg H, Bernaards C. A systematic review of correlates of sedentary behaviour in adults aged 18-65 years: a socio-ecological approach. *BMC Publ Health* 2016;**16**:163.
11. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. *Lancet Global Health* 2018;**6**(10):e1077–86.
12. Moraes A, Guerra P, Menezes P. The worldwide prevalence of insufficient physical activity in adolescents; a systematic review. *Nutr Hosp* 2013;**28**:575–84.
13. Guthold R, Stevens GA, Riley LM, Bull FC. Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1·6 million participants. *Lancet Child Adolesc Health* 2020;**4**(1):23–35.
14. Bauman A, Ainsworth B, Sallis J, Hagstromer M, Craig CL, Bull FC. The descriptive epidemiology of sitting: a 20-country comparison using the international physical activity Questionnaire (IPAQ). *Am J Prev Med* 2011;**41**:228–35.
15. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE. Sedentary behavior research network (SBRN) – terminology consensus project process and outcome. *Int J Behav Nutr Phys Activ* 2017;**14**(1):75.
16. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. *Lancet Global Health* 2018;**6**(10):e1077–86.
17. Biswas A, Oh P, Faulkner G, Bajaj RR, Silver MA, Mitchell MS. Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults A systematic review and meta-analysis. *Ann Intern Med* 2015;**162**:123–32.
18. Ekelund U, Steene-Johannessen J, Brown W, Fagerland MW, Owen N, Powell KE. Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. *Lancet* 2016;**388**(10051):1302–10.
19. Li L, Moosbrugger ME. Correlations between physical activity participation and the environment in children and adolescents: a systematic review and meta-analysis using ecological frameworks. *Int J Environ Res Publ Health* 2021;**18**(17).
20. Owen N, Sugiyama T, Eakin E, Gardiner PA, Tremblay MS, Sallis JF. Adults' sedentary behavior: determinants and interventions. *Am J Prev Med* 2011;**41**:189–96.
21. Cavazzotto T, Ronque E, Júnior H, Queiroga MR, Serassuelo H. Social-ecological correlates of regular leisure-time physical activity practice among adults. *Int J Environ Res Publ Health* 2020;**17**:3619.
22. Lin C-Y, Koohsari J, Liao Y, Ishii K, Shibata A, Nakaya T. Workplace neighbourhood built environment and workers' physically-active and sedentary behaviour: a systematic review of observational studies. *Int J Behav Nutr Phys Activ* 2020;**17**.
23. Koohsari J, Sugiyama T, Sahlqvist S, Mavoa S, Hadgraft N, Owen N. Neighborhood environmental attributes and adults' sedentary behaviors: review and research agenda. *Prev Med* 2015;**77**.
24. Yu Y-C, Ting-Fu L, Lin C-Y, Hsueh MC, Liao Y, Yang LT. Associations of the audited residential neighborhood built-environment attributes with objectively-measured sedentary time among adults: a systematic review. *Int J Environ Health Res* 2022:1–15.
25. Gavand K, Cain K, Conway T, Saelens BE, Frank LD, Kerr J. Associations between neighborhood recreation environments and adolescent physical activity. *J Phys Activ Health* 2019;**16**:1–6.
26. Parajára M, de Castro B, Coelho D, Meireles A. Are neighborhood characteristics associated with sedentary behavior in adolescents? A systematic review. *Int J Environ Health Res* 2019;**30**:1–21.
27. Cameron A, van Stralen M, Kunst A, te Velde S, van Lenthe F, Salmon J. Macroenvironmental factors including GDP per capita and physical activity in europe. *Med Sci Sports Exerc* 2012;**45**.
28. Humpel N, Owen N, Leslie E. Environmental factors associated with adults' participation in physical activity. *Am J Prev Med* 2002;**22**:188–99.
29. Ng SW, Norton E, Popkin B. Why have physical activity levels declined among Chinese adults? Findings from the 1991–2006 China health and nutrition surveys. *Soc Sci Med* 2009;**68**:1305–14. 1982.
30. EEA. *European environment–state and outlook 2015: assessment of global megatrends*. Denmark: European Environment Agency Copenhagen; 2015.
31. Dancause KN, Dehuff C, Soloway LE, Vilar M, Chan C, Wilson M. Behavioral changes associated with economic development in the South Pacific: health transition in Vanuatu. *Am J Hum Biol* 2011;**23**(3):366–76.
32. Page M, McKenzie J, Bossuyt P, Boutron I, Hoffmann T, Mulrow C. Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. *J Clin Epidemiol* 2021:134.
33. Bosdriesz JR, Witvliet MI, Visscher TL, Kunst AE. The influence of the macro-environment on physical activity: a multilevel analysis of 38 countries worldwide. *Int J Behav Nutr Phys Activ* 2012;**9**:110.
34. Van Tuyckom C. Macro-environmental factors associated with leisure-time physical activity: a cross-national analysis of EU countries. *Scand J Publ Health* 2011;**39**(4):419–26.
35. Haase A, Steptoe A, Sallis JF, Wardle J. Leisure-time physical activity in university students from 23 countries: associations with health beliefs, risk awareness, and national economic development. *Prev Med* 2004;**39**(1):182–90.
36. Wang M, Wen X, Zhang Y, Jiang C, Wang F. Is economic environment associated with the physical activity levels and obesity in Chinese adults? A cross-sectional study of 30 regions in China. *BMC Publ Health* 2017;**17**(1):701.
37. Atkinson K, Lowe S, Moore S. Human development, occupational structure and physical inactivity among 47 low and middle income countries. *Prev Med Rep* 2016;**3**:40–5.

38. Ruseski JE, Maresova K. Economic freedom, sport policy, and individual participation in physical activity: an international comparison. *Contemp Econ Pol* 2014;**32**(1):42–55.
39. Ma C, Zhang Y, Zhao M, Bovet P, Xi B. Physical activity and sedentary behavior among young adolescents in 68 lmics, and their relationships with national economic development. *Int J Environ Res Publ Health* 2020;**17**(21):1–18.
40. Linando AY, Lee S-K. Socioeconomic, dietary, nutrition, body weight and epidemiologic transitions: three nations at different stage of development. *Prog Nutr* 2018;**20**(4):602–15.
41. Van Cauwenberg J, Loyen A, Lakerveld J, Cardon G, De Craemer M, Gheysen F. Differential influences of population densification and economic growth on Europeans' physical activity and sitting time. *Cities* 2018;**82**:141–9.
42. Lee SK, Sobal J. Socio-economic, dietary, activity, nutrition and body weight transitions in South Korea. *Publ Health Nutr* 2003;**6**(7):665–74.
43. Chen X, Lau M, Kan MY, Chiang IC, HU YJ, Gong J. Lifestyle and addictive behaviors among Chinese adolescents in Hong Kong, Macau, taipei, wuhan, and zhuhai-a first cross-subculture assessment. *Int J Behav Med* 2016;**23**(5):561–70.
44. Pascual C, Regidor E, Astasio P, Ortega P, Navarro P, Domínguez V. The association of current and sustained area-based adverse socioeconomic environment with physical inactivity. *Soc Sci Med* 2007;**65**(3):454–66.
45. Wicker P, Downward P, Lera-Lopez F. Does regional disadvantage affect health-related sport and physical activity level? A multi-level analysis of individual behaviour. *Eur J Sport Sci* 2017;**17**(10):1350–9.
46. Werneck AO, Sadarangani KP, Ramírez-Vélez R, Baldew SS, Gomes TN, Ferrari G. Macroeconomic, demographic and human developmental correlates of physical activity and sitting time among South American adults. *Int J Behav Nutr Phys Activ* 2020;**17**(1):163.
47. Yang Y, Hong X, Gurney JG, Wang Y. Active travel to and from school among school-age children during 1997-2011 and associated factors in China. *J Phys Activ Health* 2017;**14**(9):684–91.
48. Schuele SA, Hilz LK, Dreger S, Bolte G. Social inequalities in environmental resources of green and blue spaces: a review of evidence in the WHO European region. *Int J Environ Res Publ Health* 2019;**16**(7).
49. Cerin ECT, Adams MA, Barnett A, Cain KL, Owen N, et al. Objectively-assessed neighbourhood destination accessibility and physical activity in adults from 10 countries: an analys is of moderators and perceptions as mediators. *Soc Sci Med* 2018;**211**:282–93.
50. Manta SW, Del Duca GF, da Silva KS, Rech CR, Gomes RdS, Maximiano GP. Is the availability of open public spaces associated with leisure-time physical activity in Brazilian adults? *Health Promot Int* 2020;**35**(1):e51–8.
51. Schipperijn J, Cerin E, Adams MA, Reis R, Smith G, Cain K. Access to parks and physical activity: an eight country comparison. *Urban For Urban Green* 2017;**27**:253–63.
52. Ong H, Gipson JD, Yang F, Li J, Li P, Liu W. The implications of rapid economic growth for health-related behaviors and chronic metabolic diseases: an exploratory study in post-earthquake west China. *J Glob Health Perspect* 2013;**2013**.
53. Rezende L, Lopes M, Rey López J, Matsudo V, Luiz O. Sedentary behavior and health outcomes: an overview of systematic reviews. *PLoS One* 2014;**9**:e105620.
54. Stevenson M, Thompson J, de Sá TH, Ewing R, Mohan D, McClure R. Land use, transport, and population health: estimating the health benefits of compact cities. *Lancet (London, England)* 2016;**388**(10062):2925–35.
55. Spence J, Rhodes R, Carson V. Challenging the dual-hinge approach to intervening on sedentary behavior. *Am J Prev Med* 2016;**52**.
56. Ranasinghe P, Pokhrel S, Anokye N. Economics of physical activity in low-income and middle- income countries: a systematic review. *BMJ Open* 2021;**11**:e037784.
57. Sedgwick P. Ecological studies: advantages and disadvantages. *BMJ Br Med J (Clin Res Ed)* 2014;**348**.
58. Ara I, Aparicio-Ugarriza R, Morales-Barco D, Nascimento de Souza W, Mata E, Gonzalez-Gross M. Physical activity assessment in the general population; validated self-report methods. *Nutr Hosp* 2015;**31**(Suppl 3):211–8.

Appendix A. All possible combinations of keywords in Search Algorithm

(a) “economic development”, “economic growth”, “economic increase”, “economic gain”, “economic expansion”, “economic depression”, “economic activation”, “economy growth”, “economy development”, “economy increase”, “economy gain”, “economy expansion”, “economy depression”, “economy activation”, “economically grow”, “economically development”, “development of economy”, “economic under development”, “economic performance”, “GDP”, “gross domestic product”, “GDP/capita”, “GDP per capita”, “GNP”, “gross national product”, “gross national income”, “national income”, “income per capita”, “TFP”, “Total Factor Productivity”; and (b) “motor activity”, “motor activities”, “sport”, “sports”, “physical fitness”, “physical exertion”, “physical activity”, “physical activities”, “physical inactivity”, “sedentary behavior”, “sedentary behavior”, “sedentary behaviors”, “sedentary behaviours”, “sedentary lifestyle”, “sedentary lifestyles”, “inactive lifestyle”, “inactive lifestyles”, “exercise”, “exercises”, “active living”, “active lifestyle”, “active lifestyles”, “outdoor activity”, “outdoor activities”, “walk”, “walking”, “running”, “bike”, “biking”, “bicycle”, “bicycling”, “cycling”, “stroll”, “strolling”, “active transport”, “ active transportation”, “active transit”, “active commuting”, “travel mode”, “physically active”, “physically inactive”, “home activity”, “home activities”, “occupational activity”, “occupational activities”, “leisure activity”, “leisure activities”, “transportation activity”, “transportation activities”, “active travel”.

Appendix B. Search Algorithm in PubMed

("economic development" [MeSH] OR "economic development" [TIAB] OR "economic growth" [TIAB] OR "economic increase" [TIAB] OR "economic gain" [TIAB] OR "economic expansion" [TIAB] OR "economic depression" [TIAB] OR "economic activation" [TIAB] OR "economy growth" [TIAB] OR "economy development" [TIAB] OR "economy increase" [TIAB] OR "economy gain" [TIAB] OR "economy expansion" [TIAB] OR "economy depression" [TIAB] OR "economy activation" [TIAB] OR "economically grow" [TIAB] OR "economically development" [TIAB] OR "development of economy" [TIAB] OR "economic under development" [TIAB] OR "economic performance" [TIAB] OR "GDP" [TIAB] OR "gross domestic product" [TIAB] OR "GDP/capita" [TIAB] OR "GDP per capita" [TIAB] OR "GNP" [TIAB] OR "gross national product" [TIAB] OR "gross national income" [TIAB] OR "national income" [TIAB] OR "income per capita" [TIAB] OR "TFP" [TIAB] OR "Total Factor Productivity" [TIAB]) AND ("exercise" [MeSH] OR "motor activity" [TIAB] OR "motor activities" [TIAB] OR "sport" [TIAB] OR "sports" [TIAB] OR "physical fitness" [TIAB] OR "physical exertion" [TIAB] OR "physical activity" [TIAB] OR "physical activities" [TIAB] OR "physical inactivity" [TIAB] OR "sedentary behavior" [TIAB] OR "sedentary behaviour" [TIAB] OR "sedentary behaviors" [TIAB] OR "sedentary behaviours" [TIAB] OR "sedentary lifestyle" [TIAB] OR "sedentary lifestyles" [TIAB] OR "inactive lifestyle" [TIAB] OR "inactive lifestyles" [TIAB] OR "exercise" [TIAB] OR "exercises" [TIAB] OR "active living" [TIAB] OR "active lifestyle" [TIAB] OR "active lifestyles" [TIAB] OR "outdoor activity" [TIAB] OR "outdoor activities" [TIAB] OR "walk" [TIAB] OR "walking" [TIAB] OR "running" [TIAB] OR "bike" [TIAB] OR "biking" [TIAB] OR "bicycle" [TIAB] OR "bicycling" [TIAB] OR "cycling" [TIAB] OR "stroll" [TIAB] OR "strolling" [TIAB] OR "active transport" [TIAB] OR "active transportation" [TIAB] OR "active transit" [TIAB] OR "active commuting" [TIAB] OR "travel mode" [TIAB] OR "physically active" [TIAB] OR "physically inactive" [TIAB] OR "home activity" [TIAB] OR "home activities" [TIAB] OR "occupational activity" [TIAB] OR "occupational activities" [TIAB] OR "leisure activity" [TIAB] OR "leisure activities" [TIAB] OR "transportation activity" [TIAB] OR "transportation activities" [TIAB] OR "active travel" [TIAB])

Public Health 215 (2023) 66–74

Contents lists available at ScienceDirect

Public Health

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

Review Paper

Insomnia and multiple health outcomes: umbrella review of meta-analyses of prospective cohort studies

T.T. Wu [a, b, c], Y.L. Zou [a, b, c], K.D. Xu [a, b, c], X.R. Jiang [a, b, c], M.M. Zhou [a, b, c], S.B. Zhang [a, b, c], C.H. Song [a, b, c, *]

[a] Department of Epidemiology and Statistics, College of Public Health, Zhengzhou University, Zhengzhou, 450001, Henan, China
[b] Henan Key Laboratory of Tumor Epidemiology, Zhengzhou University, Zhengzhou, 450052, Henan, China
[c] State Key Laboratory of Esophageal Cancer Prevention & Treatment, Zhengzhou University, Zhengzhou, 450052, Henan, China

ARTICLE INFO

Article history:
Received 30 May 2022
Received in revised form
17 September 2022
Accepted 28 November 2022
Available online 14 January 2023

Keywords:
Insomnia
Health
Umbrella review
Meta-analysis
Cohort studies

ABSTRACT

Objectives: To evaluate existing evidence of prospective cohort studies on associations between insomnia and multiple health outcomes.
Study design: An umbrella review of meta-analyses of prospective cohort studies.
Methods: A systematic search was undertaken in Pubmed, Embase, Cochrane, and Web of Science from inception to October 2021 to find meta-analyses of prospective cohort studies investigating the association of insomnia with any health outcome. The summary relative risk (SRR) for each meta-analysis was recalculated with random-effects model. The methodological quality and the quality of evidence were assessed by the A Measurement Tool to Assess Systematic Reviews and Grading of Recommendations, Assessment, Development and Evaluation, respectively.
Results: A total of 25 published meta-analyses of prospective cohort studies, reporting 63 SRRs for 29 unique outcomes were included. Insomnia was mainly related to cardiovascular outcomes and mental disorders. The former comprised atrial fibrillation (SRR: 1.30, 95% confidence interval: 1.26 to 1.35), cardiovascular diseases (1.45, 1.29 to 1.64), coronary heart disease (1.28, 1.10 to 1.50), myocardial infarction (1.42, 1.17 to 1.72), and stroke (1.55, 1.39 to 1.72). The latter involved alcohol abuse (1.35, 1.08 to 1.67), all mental disorders (2.16, 1.70 to 3.97), anxiety (3.23, 1.52 to 6.85), depression (2.31, 1.90 to 2.81), suicidal ideation (2.26, 1.79 to 2.86), suicidal attempt (1.99, 1.31 to 3.02), and suicidal death (1.72, 1.42 to 2.08). Besides, insomnia enhanced the risk of Alzheimer's disease (1.51, 1.06 to 2.14) and hyperlipidemia (1.64, 1.53 to 1.76).
Conclusion: Insomnia exhibits considerable adverse outcomes, primarily comprises cardiovascular outcomes and mental disorders, but further studies with robustly designed trials are needed to draw firmer conclusions.

Introduction

Insomnia is the most common one of seven major categories of sleep disorders,[1] characterized by a feeling of difficulty initiating sleep (DIS), difficulty maintaining sleep (DMS), early-morning awakening, or non-restorative sleep (NRS).[2] A recent meta-analysis has depicted a pooled prevalence of insomnia in 22.0% of the general population.[3] As an imperative public health challenge, insomnia pronounced effects on daytime performance, physical health, and quality of life.

Emerging evidence has indicated that insomnia may enhance risks of multiple health outcomes, such as breast cancer,[4] lung cancer,[5] hypertension,[6] peptic ulcer disease,[7] cardiovascular diseases (CVDs),[8] Alzheimer's disease (AD),[9] suicide risk,[10] depression,[11] type two diabetes mellitus (T2DM)[12] and the like. While other researchers found no meaningful associations of insomnia with breast cancer,[13] lung cancer,[14] and metabolic disruption.[15]

Contradictory results in several health outcomes were observed in existing epidemiological studies. Although abundant meta-analyses have been conducted, the results were less convincing due to different study types, different combined effect models, and

* Corresponding author. Department of Epidemiology and Statistics, College of Public Health, Zhengzhou University, Zhengzhou, Henan, 450001, China. Tel.: +86 0371 67781454.
E-mail address: sch16@zzu.edu.cn (C.H. Song).

https://doi.org/10.1016/j.puhe.2022.11.021
0033-3506/

different confounding biases. To date, no comprehensive umbrella review of the relations between insomnia and multiple health outcomes was exhibited. Thus, we conducted an umbrella review of meta-analyses to reanalyze the existing evidence of prospective cohort studies on insomnia and its effects on multiple health outcomes and to evaluate the strength and validity of these relationships. We hope to provide a theoretical basis for preventing the current public health problems affected by insomnia, and arouse people's awareness of improving sleep hygiene.

Methods

Umbrella review methods

Umbrella review is also called an overview of systematic reviews, which systematically searches, organizes, and evaluates existing evidence from multiple systematic reviews and meta-analyses on all health outcomes associated with a particular exposure.[16] This umbrella review of meta-analyses was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines,[17] and the protocol was registered with International prospective register of systematic reviews (Registration No CRD42022358424).

Literature search

Two investigators performed independent searches in Pubmed, Embase, Cochrane Database of Systematic Reviews, and Web of Science from inception to 10 October 2021 to find meta-analyses of prospective cohort studies investigating the connection between insomnia and any health outcome, without any limitation of language or publication date. Key search terms were related to insomnia ('Sleep Initiation and Maintenance Disorders' OR 'Early Awakening' OR 'Sleeplessness' and so on) and meta-analysis study design ('Systematic Review' OR 'Meta-Analysis'), the full search strategy including Medical Subject Headings terms is provided in Supplemental Table 1. We also performed a manual search of the references of eligible articles. Any discrepancies were resolved by a third investigator for the final decision. Endnote X9 was used to help in article management and selection.

Selection of meta-analyses

Studies were included if they met the following criteria: (1) investigated the association of insomnia with any health outcome; (2) included meta-analyses of prospective cohort studies that were conducted in the general population; (3) reported multivariate adjusted summary risk estimates and their corresponding 95% confidence intervals (CIs).

Meta-analyses of observational studies that combined prospective, retrospective, and cross-sectional studies in their analyses were also eligible. We excluded primary studies with the following criteria from each selected meta-analysis: retrospective cohort studies, cross-sectional studies, case–control studies, studies with unadjusted risk estimates, studies with effect sizes unable to be transformed to risk estimates, or studies conducted in diseased populations. Besides, we excluded meta-analyses that missed summary risk estimates or other important data, such as I^2 and Egger's test.

If more than one health outcome were examined in a single article, each of these health outcomes was embodied separately. If two or more meta-analyses were available for the same disease outcome, we extracted both of them while they covered different primary studies and only had little overlap. However, if the larger meta-analysis completely covered others, or there was a lot of overlap between several meta-analyses, only the largest meta-analysis with the most amount of primary studies was selected in our analysis. Any disagreement over the eligibility of particular studies was resolved through discussion with a third reviewer.

Data extraction

Two investigators independently extracted the following data from eligible meta-analyses: the first author's name, journal, year of publication, population, outcome, number of studies, adjusted risk estimates and corresponding 95% CIs, type of effect model used in the meta-analysis (fixed or random), the number of participants and events, publication bias, information on funding, and conflict of interest. Disagreements were resolved by consensus with a third reviewer.

Assessment of methodological quality

The methodological quality of each eligible meta-analysis was evaluated by A Measurement Tool to Assess Systematic Reviews (AMSTAR).[18] This reliable and valid tool contains 11 judgmental items, assessing the methodological quality of meta-analysis from the following aspects: search, analysis, and transparency. After a judgment for 11 items, each meta-analysis owns an overall score ranging from 0 to 11. Over eight points were defined as high quality, 4–7 points and three points or less were considered moderate and low quality, respectively.

Evaluation of quality of evidence

The quality of evidence was estimated using the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation).[19] GRADE rates the certainty of evidence as high, moderate, low, or very low. Observational studies start as low-certainty evidence, then can be downgraded based on the following five criteria: risk of bias, inconsistency, indirectness, imprecision, and publication bias. The risk of bias for primary studies was evaluated using the Newcastle–Ottawa scale (NOS), we assigned 'serious' when primary studies with the NOS score <7 comprised a large proportion of the published meta-analysis. There are also three criteria for upgrading: a large magnitude of association, a dose–response gradient, and attenuation by plausible confounding. We used an online tool GRADE Profiler (GRADEpro) GDT (https://gradepro.org/) to appraise the quality of evidence for each outcome explored in the umbrella review. Guideline Development Tool (GDT) is an online tool based on the GRADEpro software extension launched by the GRADE working group in 2013. After selecting the research type, five downgraded conditions, and three upgraded conditions, GDT will automatically generate the grade of evidence quality and create a table concisely summarizing the information.

Data synthesis and analysis

For each selected meta-analysis, we extracted adjusted relative risks and its 95% CIs from each primary prospective cohort study. Then, we recalculated summary relative risks (SRRs) and corresponding 95% CIs by using the DerSimonian and Laird random effects model,[20] which takes into account variance both within and between studies.

If the type of the metric was standardized mean difference, we transformed it to risk estimate before using it in the analysis. If there was insufficient data to recalculate, we directly adopted published SRR as extracted from meta-analysis. When the primary study reported risk estimates from the same cohort separately by

sex or race, we first combined the risk estimates per cohort using fixed effect models, before conducting the overall meta-analysis.

For each meta-analysis, we evaluated between-study heterogeneity by using the I^2 statistic. Whereas I^2 is dependent on the study size, we also calculated τ^2, which is independent on study size. Publication bias was assessed with the use of Egger's test. A P value less than 0.10 was taken as statistical evidence of the presence of potential publication bias. All analyses were conducted with Stata statistical analysis software version 16 (Stata Corp, College Station, Texas), with statistical significance defined as $P \leq 0.05$.

Results

A total of 4983 articles were identified by searching four electronic databases. We reviewed the titles and abstracts of all studies, and as a result, 74 articles remained. After retrieving 74 full-text articles, 25 articles were considered eligible for this umbrella review.[21–45] Supplementary Table 2 provides a list of excluded studies, along with the reasons for exclusion, after screening the full text. The flow chart illustrating the study selection process is depicted in Fig. 1.

Characteristics of included meta-analyses

The characteristics of these 25 articles are shown in Supplemental Table 3. We identified 25 published meta-analyses of prospective cohort studies, reporting 63 SRRs for 29 unique outcomes (Figs. 2 and 3). Five of these published meta-analyses reported SRRs for both insomnia and different symptom categories of insomnia (DIS, DMS, EMA, NRS) (Fig. 4). We classified 29 different outcomes into seven categories, and the map of diverse health outcomes related to insomnia was presented in Supplemental Fig. 1.

The systematic search identified four published meta-analyses for hypertension and depression, and two meta-analyses for cardio-cerebral vascular disease (CCVD), CVD mortality, suicidal ideation, suicidal attempt, diabetes mellitus, and all-cause mortality. For the other 21 outcomes, only one published meta-analysis was available.

All meta-analyses were published between 2011 and 2021 and included 2–29 primary studies. Primary studies of retrospective cohort, case–control, and cross sectional were excluded from the meta-analyses on cancer,[21] atrial fibrillation,[23] hypertension,[28,29] diabetes mellitus,[29] hyperlipidemia,[29] cognitive decline/AD,[33] suicidal ideation,[39] suicidal attempt,[39] gestational diabetes,[42] and back pain.[45] Four primary studies only reported unadjusted risk estimates and were, therefore, excluded from the meta-analyses on cancer,[21] hypertension,[30] depression,[38] and cognitive decline/AD,[33] which did not affect the results.

Cardiovascular outcomes

The recalculated SRRs with their corresponding 95% CIs and the quality of evidence for the associations between insomnia and

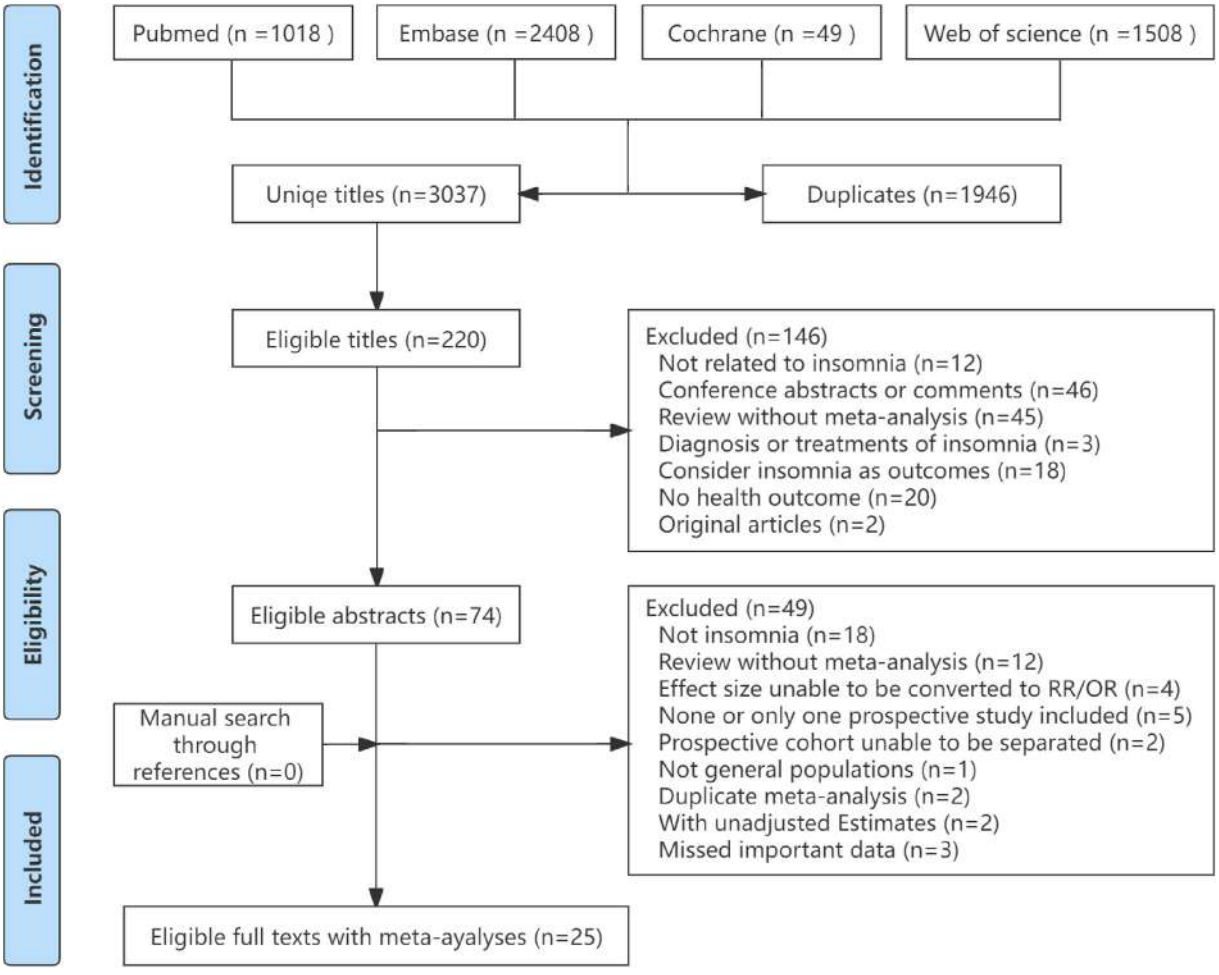


Fig. 1. Flow chart illustrating the literature search process in the umbrella review.

Outcome	No of studies	No of events/total	Recalculated SRR (95% CI)	Recalculated SRR (95% CI)	I^2 (%)	τ^2	Egger's *P*-value	AMSTAR	GRADE
Cancer outcomes									
Cancer[21]	6	11613/642644	1.14 (0.98, 1.32)		66.83	0.019	0.016	10	Very low
Cancer mortality[22]	2	5978/309121	0.94 (0.58, 1.54)		30.95	0.067	0.269	11	Moderate
Cardiovascular outcomes									
Atrial fibrillation[23]	2	NP/14161212	1.30 (1.26, 1.35)		0.00	0.000	0.417	9	Moderate
CCVD[24]	15	NP/267756	1.13 (1.07, 1.20)		53.97	0.014	0.240	9	Very low
CCVD[25]	10	NP/619593	1.13 (1.07, 1.18)		64.64	0.005	0.938	9	Very low
CVD[26]	14	6332/122501	1.45 (1.29, 1.64)		28.02	0.014	0.024	7	Very low
CVD mortality[22]	8	8357/86537	1.14 (1.03, 1.27)		20.45	0.007	0.909	11	Low
CVD mortality[27]	11	NP/212866	1.33 (1.13, 1.57)		73.08	0.054	0.002	5	Very low
CHD[27]	3	NP/150802	1.28 (1.10, 1.50)		31.30	0.008	0.100	5	Low
Hypertension[28]	2	NP/6735	1.31 (0.99, 1.72)		89.13	0.035	0.319	7	Very low
Hypertension[29]	5	30278/117829	1.87 (1.33, 2.63)		98.20	0.116	0.371	10	Very low
Hypertension[30]	13	NP/394920	1.22 (1.10, 1.34)		96.08	0.023	0.924	9	Very low
Hypertension[31]	5	NP/33003	1.18 (1.10, 1.28)		67.71	0.010	0.007	9	Very low
MI[27]	5	NP/73523	1.42 (1.17, 1.72)		48.72	0.022	0.107	5	Low
Stroke[27]	2	NP/23312	1.55 (1.39, 1.72)		0.00	0.000	0.650	5	Low
Cognitive disorders									
All-cause dementia[32]	12	23436/226167	1.17 (0.95, 1.43)		84.77	0.079	0.719	9	Very low
AD[32]	3	22079/181658	1.51 (1.06, 2.14)		57.37	0.054	0.035	9	Very low
Cognitive decline/AD[33]	6	NP/9206	1.35 (1.07, 1.71)		47.05	0.034	0.007	8	Low
Cognitive disorders[34*]	23	30027/260915	1.27 (1.16, 1.39)		82.00	NP	0.183	8	Low
Vascular dementia[32]	3	22147/183298	1.13 (0.94, 1.35)		0.00	0.000	0.662	9	Low

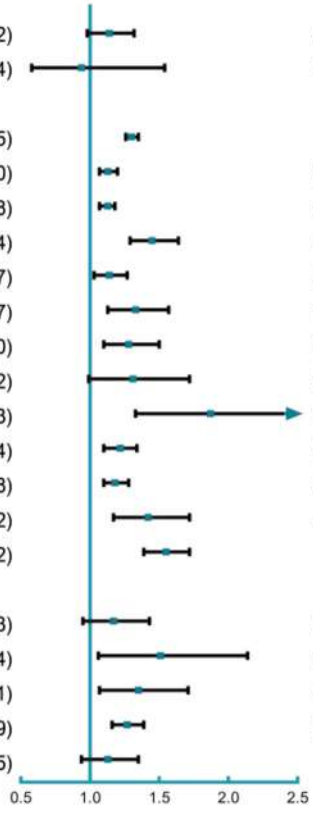


Fig. 2. Recalculated SRRs with 95% CI and quality of evidence for associations between insomnia and cancer outcomes, cardiovascular outcomes, and cognitive disorders. *Summary relative risk extracted from published meta-analysis, no reanalysis possible. SRR=Summary relative risk; CI = confidence intervals; NP = not published; CCVD=Cardio-cerebral vascular disease; CVD=Cardiovascular disease; CHD=Coronary heart disease; MI = Myocardial infarction; AD = Alzheimer's disease.

cardiovascular outcomes were shown in Fig. 2. No relationship with high quality of evidence was found. Insomnia was related to a 30% increased risk of atrial fibrillation with moderate quality of evidence. Compared with non-insomnia, risks were elevated by 45% (SRR: 1.45, 95% CI: 1.29 to 1.64) for CVD, 28% (1.28, 1.10 to 1.50) for CHD, 42% (1.42, 1.17 to 1.72) for MI, and 55% (1.55, 1.39 to 1.72) for stroke. The quality of evidence for the above links was graded as low to very low.

Two published meta-analyses indicated consistent weak relations between insomnia and CCVD with the same raised risk of 13% and consistent weak links between NRS and CCVD with similarly increased risks of 13% and 15%. In terms of EMA, one published meta-analyses represented an infinitesimally higher risk of 7%, while another showed non-significant relevance. Besides, a 26% increased risk between DIS and CCVD was observed, as well as a slightly elevated risk of 10% between DMS and CCVD.

Another two published meta-analyses revealed marginally elevated risks of CVD mortality by 14% (1.14, 1.03 to 1.27) and 33% (1.33, 1.13 to 1.57), respectively. Additionally, we also found the elevated risks of CVD mortality for DIS (1.20, 1.01, 1.43) and NRS (1.34, 1.16 to 1.54), but non-significant risk for DMS and EMA.

Four published meta-analyses represented relations between insomnia and hypertension with very low quality of evidence. Three of them produced raised risks of 87% (1.87, 1.33 to 2.63), 22% (1.22, 1.10 to 1.34), and 18% (1.18, 1.10 to 1.28), respectively, and another one was also relevant with an elevated risk but without significance. One of them showed the associations of different symptom categories, with low or very low quality of evidence. With regard to DIS, a 23% increased risk of hypertension was observed. In addition, we found elevated risks of 23% between DMS and hypertension, 13% between EMA and hypertension.

Mental disorders

Fig. 3 depicted reanalyzed SRRs with their corresponding 95% CIs and the quality of evidence for the associations between insomnia and mental disorders. None of these relationships was graded with high quality evidence. Compared with non-insomnia, risks were increased by 35% (1.35, 1.08 to 1.67) for alcohol abuse, 160% (2.16, 1.70 to 3.97) for all mental disorders, and 223% (3.23, 1.52 to 6.85) for anxiety disorders, and all with low quality of evidence.

Four published meta-analyses were in agreement regarding the direction and magnitude of associations between insomnia and depression, with medium elevated risks of 58% (1.58, 1.14 to 2.18), 183% (2.83, 1.55 to 5.18), 118% (2.18, 1.88 to 2.52), and 131% (2.31, 1.90 to 2.81), respectively.

Positive relations consistently persisted in the other two published meta-analyses for insomnia with risk of suicidal ideation, with statistically significant higher risks of 30% (1.30, 1.03 to 1.63) and 126% (2.26, 1.79 to 2.86), respectively. One of the above two published meta-analyses exhibited a 99% (1.99, 1.31 to 3.02)

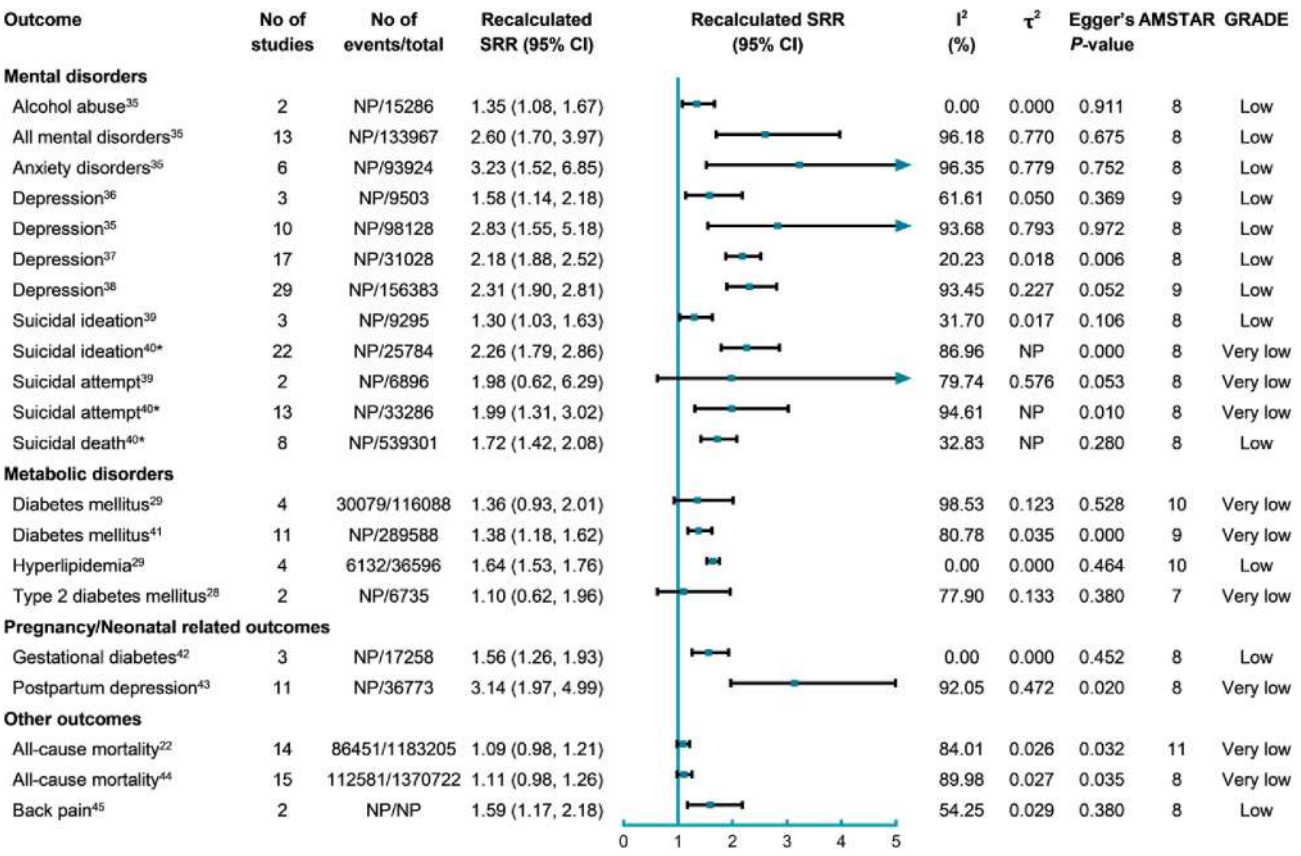


Fig. 3. Recalculated SRRs with 95% CI and quality of evidence for associations between insomnia and mental disorders, metabolic disorders, pregnancy/neonatal related outcomes, and other outcomes. *Summary relative risk extracted from published meta-analysis, no reanalysis possible. SRR=Summary relative risk; CI = confidence intervals; NP = not published.

elevated risk of insomnia and suicidal attempt, and a 72% (1.72, 1.42 to 2.08) higher risk of insomnia and suicidal death.

Cancer outcomes

The forest plot with recalculated SRRs and the quality of evidence for the associations between insomnia and cancer outcomes were exhibited in Fig. 2. Insomnia was connected with a marginally higher risk of cancer but without reaching significance and with very low quality of evidence. No significant relevance with the moderate quality of evidence was observed between insomnia and cancer mortality, as well as DIS, DMS, NRS, and cancer mortality.

Cognitive disorders

Fig. 2 showed a forest plot with the reanalyzed SRRs and the quality of evidence for the associations between insomnia and cognitive disorders. The quality of evidence for all relationships was classified as low to very low. Compared with non-insomnia, risks were increased by 51% (1.51, 1.06 to 2.14) for AD, 35% (1.35, 1.07 to 1.71) for cognitive decline/AD, and 27% (1.27, 1.16 to 1.39) for cognitive disorders, respectively. Moreover, the links between insomnia with all-cause dementia and vascular dementia showed no significance.

Metabolic disorders

The remerged SRRs with their corresponding 95% CIs and the quality of evidence for the associations between insomnia and metabolic disorders were displayed in Fig. 3. One published meta-analysis revealed elevated risks of diabetes mellitus by 38% (1.38, 1.18 to 1.62), while another presented a non-significant association, and the former also indicated increased risks of diabetes mellitus by DIS (1.55, 1.23 to 1.95) and DMS (1.72, 1.45 to 2.05), with low quality of evidence. Besides, insomnia was relevant to a higher risk of hyperlipidemia (1.64, 1.53 to 1.76). What is more, non-significant relation was observed between insomnia and T2DM.

Pregnancy/neonatal-related outcomes

Fig. 3 depicted a forest plot with recalculated SRRs and their corresponding 95% CIs, as well as the quality of evidence for the associations between insomnia and pregnancy/neonatal-related outcomes. Insomnia was related to a 56% (1.56, 1.26 to 1.93) higher risk of gestational diabetes with low quality of evidence, and a 214% (3.14, 1.97 to 4.99) increased risk of postpartum depression but with very low quality of evidence.

Other outcomes

The forest plot with recalculated SRRs and the quality of evidence for the associations between insomnia and other outcomes were exhibited in Fig. 3. Insomnia was connected with a 59% (1.59, 1.17 to 2.18) elevated risk of back pain with low quality of evidence. Two published meta-analyses consistently displayed non-significant relevance between insomnia and all-cause mortality. One of the previous two published meta-analyses presented a 10% increased risk of all-cause mortality for DIS and 23% for DMS, but

Outcome	No of studies	No of events/total	Recalculated SRR (95% CI)	Recalculated SRR (95% CI)	I^2 (%)	τ^2	Egger's *P*-value	AMSTAR	GRADE
DIS									
Cancer mortality[22]	3	1779/16925	1.21 (0.95, 1.54)		0.00	0.000	0.928	11	Low
CCVD[24]	15	NP/267756	1.26 (1.13, 1.42)		65.99	0.028	0.031	9	Very low
CVD mortality[22]	4	488650589	1.20 (1.01, 1.43)		0.00	0.000	0.484	11	Low
Hypertension[31]	4	NP/31959	1.23 (1.02, 1.47)		78.73	0.024	0.021	9	Very low
Diabetes mellitus[43]	6	NP/40649	1.55 (1.21, 1.99)		30.33	0.029	0.445	9	Low
All-cause mortality[22]	18	NP/135871	1.10 (1.00, 1.21)		70.48	0.024	0.050	11	Very low
DMS									
Cancer mortality[22]	2	1151/14394	1.11 (0.67, 1.83)		67.83	0.090	0.397	11	Low
CCVD[24]	12	NP/241760	1.10 (1.01, 1.19)		31.15	0.006	0.721	9	Low
CVD mortality[22]	3	2225/26995	1.03 (0.82, 1.30)		0.00	0.000	0.295	11	Low
Hypertension[31]	4	NP/32494	1.23 (1.06, 1.44)		75.09	0.021	0.137	9	Very low
Diabetes mellitus[43]	6	NP/40649	1.74 (1.30, 2.34)		66.93	0.079	0.610	9	Low
All-cause mortality[22]	13	9458/97843	1.04 (0.96, 1.14)		54.50	0.011	0.957	11	Very low
EMA									
CCVD[24]	9	NP/135876	1.02 (0.95, 1.10)		0.00	0.000	0.243	9	Low
CCVD[25]	4	NP/57769	1.07 (1.05,1.09)		0.32	0.000	0.338	9	Low
CVD mortality[22]	3	4258/48058	0.93 (0.76, 1.13)		0.00	0.000	0.492	11	Low
Hypertension[31]	2	NP/18408	1.13 (1.07, 1.20)		0.00	0.000	0.826	9	Low
All-cause mortality[22]	14	NP/129998	0.97 (0.91, 1.04)		38.56	0.005	0.813	11	Low
NRS									
Cancer mortality[22]	2	1689/21622	1.06 (0.53, 2.14)		83.85	0.217	0.330	11	Very low
CCVD[24]	5	NP/185145	1.17 (1.05, 1.30)		21.91	0.003	0.968	9	Low
CCVD[25]	5	NP/109751	1.15 (1.06, 1.26)		44.97	0.004	0.903	9	Low
CVD mortality[22]	3	3714/45069	1.34 (1.16, 1.54)		0.00	0.000	0.288	11	Low
All-cause mortality[22]	7	NP/69841	1.23 (1.07, 1.42)		64.96	0.020	0.035	11	Very low

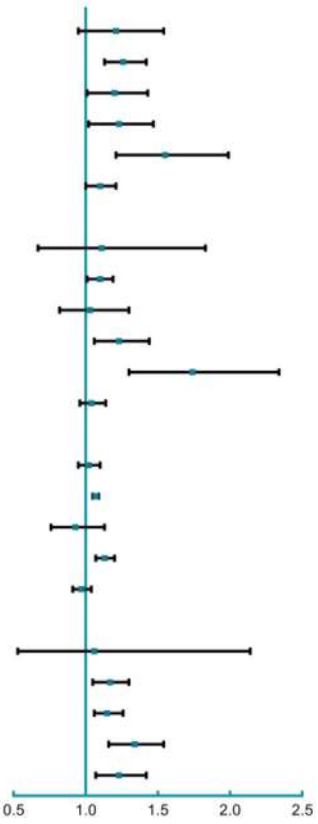


Fig. 4. Recalculated SRRs with 95% CI and quality of evidence for associations between different symptom categories (DIS, DMS, EMA, NRS) and several outcomes. *Summary relative risk extracted from published meta-analysis, no reanalysis possible. DIS = difficulty initiating sleep; DMS = difficulty maintaining sleep; EMA = early-morning awakening; NRS = non-restorative sleep; SRR=Summary relative risk; CI = confidence intervals; NP = not published.

the non-significant association of DMS and EMA with all-cause mortality.

Methodological quality

The median AMSTAR score achieved across all studies was eight out of 11 (range 5–11, interquartile range 8–9). Of the 25 eligible meta-analyses, 22 (88%) meta-analyses were conducted with a high-quality approach (≥8 points), and the other 3 (12%) were performed with a moderate quality method (4–7 points). In general, the main shortcomings of the published meta-analyses were that data extraction or study selection was unable to be duplicated, grey literature was not accounted for in the literature search, no list of excluded studies was provided, and study quality was not assessed. A breakdown of AMSTAR scores for each included meta-analysis is available in Supplemental Table 4.

Quality of evidence

Of the 63 SRRs reported in this umbrella review, the quality of the evidence was rated moderate for 3% of the associations (n = 2), low for 54% of the associations (n = 34), and very low for 43% of the associations (n = 27) with the GRADE classification. None of these relationships was rated with high quality of evidence because of that cohort studies were appraised with low-quality evidence at the beginning while lacking plausible confounding analysis and dose–response effect. Only three SRRs identified as having plausible confounding analyses and 4 identified as having a large magnitude of effects, and without other significant bias reached a GRADE classification of 'moderate' compared with the majority rating of 'low'. A breakdown of GRADE scores for eligible meta-analyses is presented in Supplemental Table 5.

Heterogeneity between primary studies

We were able to recalculate I^2 and τ^2 of each SRR with random effects model apart from 6% (n = 4) of the total 63 SRRs. For 4 SRRs that were unable to be recalculated, we directly extracted published I^2 from corresponding meta-analyses. Finally, a total of 63 SRRs presented I^2, about 56% (n = 35) had large heterogeneity ($I^2 > 50\%$), and 33% (n = 21) had very large heterogeneity ($I^2 > 75\%$).

Publication bias of included studies

We performed Egger's test in 94% (n = 59) of SRRs in our reanalysis. For 4 SRRs that were unable to be recalculated, we directly extracted reported Egger's *P*-value from corresponding meta-analyses. Finally, a total of 63 SRRs presented Egger's *P*-value. According to Egger's test ($P < 0.10$), our results indicated the presence of potential publication bias for 19 SRRs, consisting of cancer, CCVD, CVD, CVD mortality, two for hypertension, AD, cognitive decline/AD, two for depression, suicidal ideation, two for

suicidal attempt, diabetes mellitus, postpartum depression, and 4 for all-cause mortality.

Discussion

Principal findings and possible interpretation

In this umbrella review, we provided a broad overview of the existing evidence and evaluated the methodological quality of the included meta-analyses and the quality of evidence for all these links. We identified 25 meta-analyses of prospective cohort studies comprising 29 unique outcomes with 63 SRRs. Notable relations were observed between insomnia and cardiovascular outcomes and mental disorders. The methodological quality was high for most of the published meta-analyses. None of these relationships was rated with high quality evidence. And the quality of evidence was graded as moderate only for two outcomes, thus further research studies would be likely to change the overall summary estimates.

The largest proportion of the included meta-analyses was about cardiovascular outcomes. Insomnia was positively related to a wide range of CVDs covering AF, CCVD, CVD, CVD mortality, CHD, hypertension, MI, and stroke. With similar results, several lines of evidence have pointed out the adverse impact of insomnia on the cardiovascular system. A cross-sectional study found that insomnia increased the risk of AF to 1.92, a little higher than our result (1.30). Notably, their subgroup analysis exhibited an extraordinary high risk of AF (6.52) in people aged <40 years.[46] Besides, a 10-year cohort study disclosed the same effect of insomnia in CVD as the present study. Four kinds of insomnia symptoms were totally associated with an increased risk of CVD, especially among young adults or adults who have not developed hypertension.[47] Therefore, young people with insomnia are the focus and should attach more importance to cardiovascular health. The main effect of insomnia is inadequate sleep duration, which increases the risks of quite a variety of cardiovascular outcomes via several biologically plausible mechanisms. First, insomnia contributes to reduced energy expenditure and up-regulation of appetite and as a consequence, glucose intolerance, impaired glycemic control, and insulin sensitivity are altered.[48,49] Besides, insomnia may enhance the activity of the sympathetic nervous system, and thus boost the renin-angiotensin-aldosterone system to release catecholamine. Accumulation of catecholamine may prompt the constriction of blood vessels and the onset of hypertension.[50] Simultaneously, a wicked circle may be established since a stimulated sympathetic nervous system in turn aggravate insomnia.[49] In addition, magnesium, as an essential trace element for human, is considered to be a physiologic calcium antagonist playing a crucial role in decreasing vascular tension. Nevertheless, this function may be impaired due to the decreased intracellular magnesium resulting from insomnia.[51]

In terms of metabolic disorders, two meta-analyses displayed inconsistent relation between insomnia and diabetes mellitus, and both of them were rated as very low-quality evidence. A recent cross-sectional study[52] suggested that insomnia was significantly relevant to diabetes mellitus in the northern Chinese population. Hence, further study could be performed to explore the certain connection. Of note, we observed that insomnia raised the risk of hyperlipidemia, and this result could be supported by a cross-sectional survey[53] indicated that frequent insomnia was related to a higher prevalence of dyslipidemia in women.

To a large extent, mental disorders were tightly influenced by insomnia, mainly involved anxiety, depression, and suicidal-related outcomes. This result was consistent with plentiful previous studies. In 2016, an observational cross-sectional study[54] reported that anxiety was more frequent in students suffering from clinical insomnia. In 2019, a cross-sectional analysis of the Canadian Longitudinal Study illustrated that people with insomnia disorder exhibited greater proportions of adverse lifestyle features such as anxiety, depression,[55] and alleged that insomnia disorder might increase the risk of further cognitive decline, which was almost similar to our study. There were several plausible biological mechanisms underlying these associations. As we all know, as a monoamine neurotransmitter, dopamine is essential for motivation, reward processing, and the ability to experience a pleasure. Unfortunately, insomnia may reduce dopamine D2/D3 receptor availability in the striatum, thus leading to underactive reward processing and anhedonia.[56] Similarly, serotonin (5-HT) is also a monoamine neurotransmitter that is crucial for attention, cognition, information processing, and mood. Insomnia is prone to disrupt the 5-HT system, resulting in an elevated risk of anxiety, depression.[57] Finally, other proposed mechanisms included increasing levels of inflammatory markers, such as C-reactive protein and interleukin-6.[58]

Our umbrella review verified that insomnia raised the risk of AD but had no influence on vascular dementia. Lacking sleep results in amyloid-β (Aβ) formation and aggregation, which are pivotal components of AD pathology.[59] Screening insomnia may help identify individuals who are at elevated risk of developing AD and implement preventive strategies, especially in the elderly.

As regards cancer outcomes, more previous studies looked at insomnia after a diagnosis of cancer, and less considered insomnia as a risk factor for cancer among primarily cancer-free individuals.[60] Most meta-analyses demonstrated a non-significant association of insomnia with the oncogenesis and progression of the carcinoma.[49] Similarly, non-significant links between insomnia and cancer-related outcomes were observed in this umbrella review. However, before excluding one unadjusted risk estimate, we found a significant but slight relevance between insomnia and cancer (SRR: 1.20; 95% CI: 1.02 to 1.42). In more detail, the pooled hazard ratio (HR) was only significantly higher in thyroid cancer (HR: 1.36; 95% CI: 1.12 to 1.65), and not in endometrial cancer, breast cancer, lung cancer, and prostate cancer. There were only one or two primary prospective cohort studies available for the above specific types of cancer. Thus, further prospective cohort studies are needed to fully investigate the connection of insomnia with specific cancer risks.

Strengths and limitations

Our umbrella review had several strengths. First, systematic methods were employed in the present review, embodying a robust search strategy of four scientific literature databases with independent study selection and extraction by two investigators, and standard approaches to assess the methodological quality (AMSTAR) and the quality of the evidence (GRADE). Second, only prospective cohort studies and studies with adjusted risk estimates were adopted in the eligible meta-analyses, contributing to considerably less bias and confounding. Third, we repeated each meta-analysis with random effects model if possible, to recalculate SRRs and 95% CIs and reappraise heterogeneity and publication bias for better comparison across outcomes. Fourth, except for insomnia, associations of different symptom categories (DIS, DMS, EMA, NRS) with health outcomes were separately explored.

Some potential limitations could not be ignored. First of all, insomnia is a subjective disease, its diagnosis mainly bases on the self-report of patients, which leads to the lack of consistency, may exert a certain influence on the results of this study. Second, there were 65% of analyses of the associations ($n = 41$) were conducted with <10 primary prospective cohort studies; hence, the interpretation of these results might be limited by the low number of studies. Third, high quality evidence for the association of insomnia

with health outcomes could not be found. Because studies of insomnia are all observational studies, starting as low quality evidence. It is almost impossible to achieve high-quality evidence. Finally, only published SR and MA were considered, and some vital recently published individual studies might be omitted in this umbrella review; thus, we considered as much of the latest research as possible in the discussion section.

Conclusion

In conclusion, insomnia exhibits considerable adverse outcomes, mainly cardiovascular outcomes and mental disorders. Lacking high-quality evidence, further studies, and robustly designed trials are needed to clarify the effect of insomnia on multiple aspects of human health and to draw firmer conclusions.

Author statements

Acknowledgements

We thank all collaborators for their support and help in carrying out this study.

Ethical approval

None declared.

Funding

This work was supported by the Henan University Science and Technology Innovation Talents Support Program [Grant 19HASTIT005]; the National Key Research and Development Program [Grant 2017YFC1309200]; and the Medical Science and Technology Key Projects of Henan Province and Zhengzhou [Grant numbers 192102310088 and 19A32000820].

Competing interests

The authors declare no competing interests.

Authors' contributions

Idea and design: WTT, JXR. Literature search: ZYL, XKD. Data extraction and analysis: WTT, ZMM, ZSB. Manuscript writing: WTT. Manuscript revision: SCH, WTT. All authors read and approved the version of the manuscript to be published.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2022.11.021.

References

1. Sateia MJ. International classification of sleep disorders-third edition: highlights and modifications. *Chest* 2014;**146**(5):1387–94. https://doi.org/10.1378/chest.14-0970.
2. Buysse DJ, Rush AJ, Reynolds CF. Clinical management of insomnia disorder. *JAMA* 2017;**318**(20):1973–4. https://doi.org/10.1001/jama.2017.15683.
3. Zeng LN, Zong QQ, Yang Y, Zhang L, Xiang YF, Ng CH, et al. Gender difference in the prevalence of insomnia: a meta-analysis of observational studies. *Front Psychiatr* 2020;**11**:577429. https://doi.org/10.3389/fpsyt.2020.577429.
4. Liu HP, Wei JC, Yip HT, Yeh MH. Association of insomnia, depressive disorders, and mood disorders as risk factors with breast cancer: a nationwide population-based cohort study of 232,108 women in taiwan. *Front Oncol* 2021;**11**:757626. https://doi.org/10.3389/fonc.2021.757626.
5. Huo Z, Ge F, Li C, Cheng H, Lu Y, Wang R, et al. Genetically predicted insomnia and lung cancer risk: a Mendelian randomization study. *Sleep Med* 2021;**87**:183–90. https://doi.org/10.1016/j.sleep.2021.06.044.
6. Lunyera J, Park YM, Ward JB, Gaston SA, Bhavsar NA, Muntner P, et al. A prospective study of multiple sleep dimensions and hypertension risk among white, black and Hispanic/Latina women: findings from the Sister Study. *J Hypertens* 2021;**39**(11):2210–9. https://doi.org/10.1097/hjh.0000000000002929.
7. Zha LF, Dong JT, Wang JL, Chen QW, Wu JF, Zhou YC, et al. Effects of insomnia on peptic ulcer disease using mendelian randomization. *Oxid Med Cell Longev* 2021;**2021**:2216314. https://doi.org/10.1155/2021/2216314.
8. Tao F, Cao Z, Jiang Y, Fan N, Xu F, Yang H, et al. Associations of sleep duration and quality with incident cardiovascular disease, cancer, and mortality: a prospective cohort study of 407,500 UK biobank participants. *Sleep Med* 2021;**81**:401–9. https://doi.org/10.1016/j.sleep.2021.03.015.
9. Baek MS, Han K, Kwon HS, Lee YH, Cho H, Lyoo CH. Risks and prognoses of alzheimer's disease and vascular dementia in patients with insomnia: a nationwide population-based study. *Front Neurol* 2021;**12**:611446. https://doi.org/10.3389/fneur.2021.611446.
10. Anna Karin H, Hössjer O, Bellocco R, Ye W, Trolle LY, Åkerstedt T. Insomnia in the context of short sleep increases suicide risk. *Sleep* 2021;**44**(4):zsaa245. https://doi.org/10.1093/sleep/zsaa245.
11. Chen SJ, Zhang JH, Li SX, Tsang CC, Chan KCC, Au CT, et al. The trajectories and associations of eveningness and insomnia with daytime sleepiness, depression and suicidal ideation in adolescents: a 3-year longitudinal study. *J Affect Disord* 2021;**294**:533–42. https://doi.org/10.1016/j.jad.2021.07.033.
12. Yao W, Luo J, Yu X, Jiang W, Zhang D. Insomnia symptoms are associated with an increased risk of type 2 diabetes mellitus among adults aged 50 and older. *Sleep & breathing = Schlaf & Atmung* 2021. https://doi.org/10.1007/s11325-021-02497-8.
13. Richmond RC, Anderson EL, Dashti HS, Jones SE, Lane JM, Strand LB, et al. Investigating causal relations between sleep traits and risk of breast cancer in women: mendelian randomisation study. *Br Med J Int Ed* 2019;**365**:l2327. https://doi.org/10.1136/bmj.l2327.
14. Xie J, Zhu M, Ji M, Fan J, Huang Y, Wei X, et al. Relationships between sleep traits and lung cancer risk: a prospective cohort study in UK Biobank. *Sleep* 2021;**44**(9):zsab089. https://doi.org/10.1093/sleep/zsab089.
15. Bos MM, Goulding NJ, Lee MA, Hofman A, Bot M, Pool R, et al. Investigating the relationships between unfavourable habitual sleep and metabolomic traits: evidence from multi-cohort multivariable regression and Mendelian randomization analyses. *BMC Med* 2021;**19**(1):69. https://doi.org/10.1186/s12916-021-01939-0.
16. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. *Int J Evid Base Healthc* 2015;**13**(3):132–40. https://doi.org/10.1097/XEB.0000000000000055.
17. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Br Med J Int Ed* 2021;**372**:n71. https://doi.org/10.1136/bmj.n71.
18. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol* 2009;**62**(10):1013–20. https://doi.org/10.1016/j.jclinepi.2008.10.009.
19. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;**64**(4):383–94. https://doi.org/10.1016/j.jclinepi.2010.04.026.
20. DerSimonian R, Laird N. Meta-analysis in clinical trials revisited. *Contemp Clin Trials* 2015;**45**(Pt A):139–45. https://doi.org/10.1016/j.cct.2015.09.002.
21. Shi T, Min M, Sun C, Zhang Y, Liang M, Sun Y. Does insomnia predict a high risk of cancer? A systematic review and meta-analysis of cohort studies. *J Sleep Res* 2020;**29**(1):e12876. https://doi.org/10.1111/jsr.12876.
22. Ge L, Guyatt G, Tian J, Pan B, Chang Y, Chen Y, et al. Insomnia and risk of mortality from all-cause, cardiovascular disease, and cancer: systematic review and meta-analysis of prospective cohort studies. *Sleep Med Rev* 2019;**48**:101215. https://doi.org/10.1016/j.smrv.2019.101215.
23. Chokesuwattanaskul R, Thongprayoon C, Sharma K, Congrete S, Tanawuttiwat T, Cheungpasitporn W. Associations of sleep quality with incident atrial fibrillation: a meta-analysis. *Intern Med J* 2018;**48**(8):964–72. https://doi.org/10.1111/imj.13764.
24. He Q, Zhang P, Li GX, Dai HX, Shi JP. The association between insomnia symptoms and risk of cardio-cerebral vascular events: a meta-analysis of prospective cohort studies. *Eur J Prev Cardiol* 2017;**24**(10):1071–82. https://doi.org/10.1177/2047487317702043.
25. Hu S, Lan T, Wang Y, Ren L. Individual insomnia symptom and increased hazard risk of cardiocerebral vascular diseases: a meta-analysis. *Front Psychiatr* 2021;**12**:654719. https://doi.org/10.3389/fpsyt.2021.654719.
26. Sofi F, Cesari F, Casini A, Macchi C, Abbate R, Gensini GF. Insomnia and risk of cardiovascular disease: a meta-analysis. *Eur J Prev Cardiol* 2014;**21**(1):57–64. https://doi.org/10.1177/2047487312460020.
27. Li M, Zhang XW, Hou WS, Tang ZY. Insomnia and risk of cardiovascular disease: a meta-analysis of cohort studies. *Int J Cardiol* 2014;**176**(3):1044–7. https://doi.org/10.1016/j.ijcard.2014.07.284.
28. Johnson KA, Gordon CJ, Chapman JL, Hoyos CM, Marshall NS, Miller CB, et al. The association of insomnia disorder characterised by objective short sleep duration with hypertension, diabetes and body mass index: a systematic

review and meta-analysis. *Sleep Med Rev* 2021;**59**:101456. https://doi.org/10.1016/j.smrv.2021.101456.
29. Zhang Y, Jiang X, Liu J, Lang Y, Liu Y. The association between insomnia and the risk of metabolic syndrome: a systematic review and meta-analysis. *J Clin Neurosci : official journal of the Neurosurgical Society of Australasia* 2021;**89**:430–6. https://doi.org/10.1016/j.jocn.2021.05.039.
30. Li L, Gan Y, Zhou X, Jiang H, Zhao Y, Tian Q, et al. Insomnia and the risk of hypertension: a meta-analysis of prospective cohort studies. *Sleep Med Rev* 2021;**56**:101403. https://doi.org/10.1016/j.smrv.2020.101403.
31. Meng L, Zheng Y, Hui R. The relationship of sleep duration and insomnia to risk of hypertension incidence: a meta-analysis of prospective cohort studies. *Hypertens Res : official journal of the Japanese Society of Hypertension* 2013;**36**(11):985–95. https://doi.org/10.1038/hr.2013.70.
32. Shi L, Chen SJ, Ma MY, Bao YP, Han Y, Wang YM, et al. Sleep disturbances increase the risk of dementia: a systematic review and meta-analysis. *Sleep Med Rev* 2018;**40**:4–16. https://doi.org/10.1016/j.smrv.2017.06.010.
33. Bubu OM, Brannick M, Mortimer J, Umasabor-Bubu O, Se Bastiao YV, Wen Y, et al. Sleep, cognitive impairment, and alzheimer's disease: a systematic review and meta-analysis. *Sleep* 2017;**40**(1). https://doi.org/10.1093/sleep/zsw084.
34. Xu W, Tan CC, Zou JJ, Cao XP, Tan L. Sleep problems and risk of all-cause cognitive decline or dementia: an updated systematic review and meta-Analysis. *J Neurol Neurosurg Psychiatry* 2019;**91**(3):236–44. https://doi.org/10.1136/jnnp-2019-321896.
35. Hertenstein E, Feige B, Gmeiner T, Kienzler C, Spiegelhalder K, Johann A, et al. Insomnia as a predictor of mental disorders: a systematic review and meta-analysis. *Sleep Med Rev* 2019;**43**:96–105. https://doi.org/10.1016/j.smrv.2018.10.006.
36. Marino C, Andrade B, Campisi SC, Wong M, Zhao HY, Jing X, et al. Association between disturbed sleep and depression in children and youths A systematic review and meta-analysis of cohort studies. *JAMA Netw Open* 2021;**4**(3):e212373. https://doi.org/10.1001/jamanetworkopen.2021.2373.
37. Baglioni C, Battagliese G, Feige B, Spiegelhalder K, Nissen C, Voderholzer U, et al. Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies. *J Affect Disord* 2011;**135**(1–3):10–9. https://doi.org/10.1016/j.jad.2011.01.011.
38. Li L, Wu C, Gan Y, Qu X, Lu Z. Insomnia and the risk of depression: a meta-analysis of prospective cohort studies. *BMC Psychiatr* 2016;**16**(1):375. https://doi.org/10.1186/s12888-016-1075-3.
39. Liu JW, Tu YK, Lai YF, Lee HC, Tsai PS, Chen TJ, et al. Associations between sleep disturbances and suicidal ideation, plans, and attempts in adolescents: a systematic review and meta-analysis. *Sleep* 2019;**42**(6):zsz054. https://doi.org/10.1093/sleep/zsz054.
40. Liu RT, Steele SJ, Hamilton JL, Do QBP, Furbish K, Burke TA, et al. Sleep and suicide: a systematic review and meta-analysis of longitudinal studies. *Clin Psychol Rev* 2020;**81**:101895. https://doi.org/10.1016/j.cpr.2020.101895.
41. Anothaisintawee T, Reutrakul S, Van Cauter E, Thakkinstian A. Sleep disturbances compared to traditional risk factors for diabetes development: systematic review and meta-analysis. *Sleep Med Rev* 2016;**30**:11–24. https://doi.org/10.1016/j.smrv.2015.10.002.
42. Zhu B, Shi C, Park CG, Reutrakul S. Sleep quality and gestational diabetes in pregnant women: a systematic review and meta-analysis. *Sleep Med* 2020;**67**:47–55. https://doi.org/10.1016/j.sleep.2019.11.1246.
43. Maghami M, Shariatpanahi SP, Habibi D, Heidari-Beni M, Badihian N, Hosseini M, et al. Sleep disorders during pregnancy and postpartum depression: a systematic review and meta-analysis. *Int J Dev Neurosci : the official journal of the International Society for Developmental Neuroscience* 2021;**81**(6):469–78. https://doi.org/10.1002/jdn.10118.
44. Lovato N, Lack L. Insomnia and mortality: a meta-analysis. *Sleep Med Rev* 2019;**43**:71–83. https://doi.org/10.1016/j.smrv.2018.10.004.
45. Amiri S, Behnezhad S. Sleep disturbances and back pain: systematic review and meta-analysis. *Neuropsychiatry* 2020;**34**(2):74–84. https://doi.org/10.1007/s40211-020-00339-9.
46. Han X, Yang Y, Chen Y, Gao L, Yin X, Li H, et al. Association between insomnia and atrial fibrillation in a Chinese population: a cross-sectional study. *Clin Cardiol* 2017;**40**(9):765–9. https://doi.org/10.1002/clc.22731.
47. Zheng B, Yu C, Lv J, Guo Y, Bian Z, Zhou M, et al. Insomnia symptoms and risk of cardiovascular diseases among 0.5 million adults: a 10-year cohort. *Neurology* 2019;**93**(23):e2110–20. https://doi.org/10.1212/WNL.0000000000008581.
48. Wang C, Bangdiwala S, Rangarajan S, Lear S, AlHabib K, Mohan V, et al. Association of estimated sleep duration and naps with mortality and cardiovascular events: a study of 116 632 people from 21 countries. *Eur Heart J* 2019;**40**(20):1620–9. https://doi.org/10.1093/eurheartj/ehy695.
49. Li J, Cao D, Huang Y, Chen Z, Wang R, Dong Q, et al. Sleep duration and health outcomes: an umbrella review. *Sleep & breathing = Schlaf & Atmung* 2021. https://doi.org/10.1007/s11325-021-02458-1.
50. Castro-Diehl C, Diez Roux A, Redline S, Seeman T, McKinley P, Sloan R, et al. Sleep duration and quality in relation to autonomic nervous system measures: the multi-ethnic study of atherosclerosis (MESA). *Sleep* 2016;**39**(11):1927–40. https://doi.org/10.5665/sleep.6218.
51. Takase B, Akima T, Uehata A, Ohsuzu F, Ajcc Kurita. Effect of chronic stress and sleep deprivation on both flow-mediated dilation in the brachial artery and the intracellular magnesium level in humans. *Clin Cardiol* 2004;**27**(4):223–7. https://doi.org/10.1002/clc.4960270411.
52. Zhang Y, Lin Y, Zhang J, Li L, Liu X, Wang T, et al. Association between insomnia and type 2 diabetes mellitus in Han Chinese individuals in Shandong Province, China. *Sleep & breathing = Schlaf & Atmung* 2019;**23**(1):349–54. https://doi.org/10.1007/s11325-018-1687-6.
53. Zhan Y, Zhang F, Lu L, Wang J, Sun Y, Ding R, et al. Prevalence of dyslipidemia and its association with insomnia in a community based population in China. *BMC Publ Health* 2014;**14**:1050. https://doi.org/10.1186/1471-2458-14-1050.
54. Choueiry N, Salamoun T, Jabbour H, El Osta N, Hajj A, Rabbaa Khabbaz Ljpo. Insomnia and relationship with anxiety in university students: a cross-sectional designed study. *PLoS One* 2016;**11**(2):e0149643. https://doi.org/10.1371/journal.pone.0149643.
55. Cross N, Carrier J, Postuma R, Gosselin N, Kakinami L, Thompson C, et al. Association between insomnia disorder and cognitive function in middle-aged and older adults: a cross-sectional analysis of the Canadian Longitudinal Study on Aging. *Sleep* 2019;**42**(8). https://doi.org/10.1093/sleep/zsz114.
56. Volkow ND, Tomasi D, Wang G-J, Telang F, Fowler JS, Logan J, et al. Evidence that sleep deprivation downregulates dopamine D2R in ventral striatum in the human brain. *J Neurosci* 2012;**32**(19):6711–7. https://doi.org/10.1523/JNEUROSCI.0045-12.2012.
57. Blake MJ, Trinder JA, Allen NB. Mechanisms underlying the association between insomnia, anxiety, and depression in adolescence: implications for behavioral sleep interventions. *Clin Psychol Rev* 2018;**63**:25–40. https://doi.org/10.1016/j.cpr.2018.05.006.
58. Faraut B, Boudjeltia KZ, Vanhamme L, Kerkhofs M. Immune, inflammatory and cardiovascular consequences of sleep restriction and recovery. *Sleep Med Rev* 2012;**16**(2):137–49. https://doi.org/10.1016/j.smrv.2011.05.001.
59. Wang C, Holtzman DM. Bidirectional relationship between sleep and Alzheimer's disease: role of amyloid, tau, and other factors. *Neuropsychopharmacology* 2020;**45**(1):104–20. https://doi.org/10.1038/s41386-019-0478-5.
60. Liang X, Harris H, Hendryx M, Shadyab A, Hale L, Li Y, et al. Sleep characteristics and risk of ovarian cancer among postmenopausal women. *Cancer Prev Res (Philadelphia, Pa)* 2021;**14**(1):55–64. https://doi.org/10.1158/1940-6207.Capr-20-0174.

Public Health 215 (2023) 1–11

Contents lists available at ScienceDirect

Public Health

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

Original Research

Patient-reported health outcomes of SARS-CoV-2–tested patients presenting to emergency departments: a propensity score–matched prospective cohort study

R. Bola [a], J. Sutherland [b], R.A. Murphy [a, c], M. Leeies [d, e], L. Grant [f, g], J. Hayward [h], P. Archambault [i], L. Graves [j], T. Rose [j], C. Hohl [k, l, *]

[a] School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada
[b] Centre for Health Services and Policy Research, University of British Columbia, Vancouver, BC, Canada
[c] Cancer Control Research, BC Cancer, Vancouver, BC, Canada
[d] Department of Emergency Medicine, University of Manitoba, Winnipeg, MB, Canada
[e] Section of Critical Care Medicine, University of Manitoba, Winnipeg, MB, Canada
[f] Department of Emergency Medicine, McGill University, Montreal, QC, Canada
[g] Emergency Department, Jewish General Hospital, Montreal, QC, Canada
[h] Department of Emergency Medicine, University of Alberta, AB, Canada
[i] Université Laval, Department of Family Medicine and Emergency Medicine, QC, Canada
[j] Patient Partner, Canadian COVID-19 Emergency Department Rapid Response Network Patient Engagement Committee, Canada
[k] Department of Emergency Medicine, University of British Columbia, Vancouver, BC, Canada
[l] Emergency Department, Vancouver General Hospital, Vancouver, BC, Canada

ARTICLE INFO

Article history:
Received 20 July 2022
Received in revised form
21 November 2022
Accepted 23 November 2022
Available online 1 December 2022

Keywords:
CCEDRRN
COVID-19
Emergency departments
Patient-reported outcomes
Propensity scores

ABSTRACT

Objective: This study aimed to compare the long-term physical and mental health outcomes of matched severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)-positive and SARS-CoV-2–negative patients controlling for seasonal effects.
Study design: This was a retrospective cohort study.
Methods: This study enrolled patients presenting to emergency departments participating in the Canadian COVID-19 Emergency Department Rapid Response Network. We enrolled consecutive eligible consenting patients who presented between March 1, 2020, and July 14, 2021, and were tested for SARS-CoV-2. Research assistants randomly selected four site and date-matched SARS-CoV-2–negative controls for every SARS-CoV-2–positive patient and interviewed them at least 30 days after discharge. We used propensity scores to match patients by baseline characteristics and used linear regression to compare Veterans RAND 12-item physical health component score (PCS) and mental health component scores (MCS), with higher scores indicating better self-reported health.
Results: We included 1170 SARS-CoV-2–positive patients and 3716 test-negative controls. The adjusted mean difference for PCS was 0.50 (95% confidence interval [CI]: -0.36, 1.36) and -1.01 (95% CI: -1.91, -0.11) for MCS. Severe disease was strongly associated with worse PCS ($\beta = -7.4$; 95% CI: -9.8, -5.1), whereas prior mental health illness was strongly associated with worse MCS ($\beta = -5.4$; 95% CI: -6.3, -4.5).
Conclusion: Physical health, assessed by PCS, was similar between matched SARS-CoV-2–positive and SARS-CoV-2–negative patients, whereas mental health, assessed by MCS, was worse during a time when the public experienced barriers to care. These results may inform the development and prioritization of support programs for patients.

* Corresponding author. 828 W 12th Ave, Vancouver, Canada.
E-mail address: corinne.hohl@ubc.ca (C. Hohl).

https://doi.org/10.1016/j.puhe.2022.11.016

Research in context

Evidence before this study

We searched for the terms "patient-reported outcomes," "coronavirus disease," "COVID-19," "mental health," and "physical health" in PubMed and Google Scholar to identify studies examining the physical or mental health outcomes of COVID-19 survivors. Most studies examining physical and mental health outcomes were from China, the United States, and Italy. Persistent fatigue, malaise, anxiety, and depression were commonly reported symptoms after the acute infection resolved. Few studies examined patient-reported outcomes of COVID-19 survivors or compared them with test-negative controls, and none incorporated demographic or sociocultural factors such as income, race, education, or employment status.

Added value of this study

We examined a cohort of consenting site and date-matched patients who presented to Canadian Emergency Departments and were tested for SARS-CoV-2. SARS-CoV-2–positive patients reported similar physical outcomes but significantly worse mental health outcomes compared with those who tested negative for SARS-CoV-2. The risk factors for worse physical outcomes included complex comorbidities (immune disorders, neurological disease, etc.), illicit substance use, female sex, arrival by ambulance, atrial fibrillation, congestive heart failure, and severe disease at presentation. The risk factors for worse mental health outcomes included illicit substance use, female sex, lower income, unstable housing, living in a correctional facility, prior mental illness, and severe disease.

Implications of all the available evidence

This is the first study that compared physical and mental health outcomes of COVID-19 patients presenting to emergency departments (EDs) with site and date-matched test-negative controls. We identified several clinical, social, and demographic variables that were associated with worse health outcomes during the study period. Future research including complications and sequelae in their examination of physical and mental health outcomes of COVID-19 patients should be conducted to guide evidence-based policies and interventions aimed at reducing the longitudinal effects of COVID-19.

Introduction

Sequelae of COVID-19 infection have been described up to 12 months postinfection in cohorts from China, the United States, and Italy.[1–3] These studies found persistent physical and mental health symptoms, including fatigue, general weakness, depression, and anxiety among survivors.[4–6] Disease severity was a predictor of physical and mental sequelae.[5–7]

Studies that examined patient-reported quality of life among COVID-19 survivors have not compared the quality of life measures with those of patients who tested negative for SARS-CoV-2. Patient-reported outcomes represent a patient's perspective of their own health and well-being and are central to their experience of illness and recovery.[8] Likewise, current COVID-19 research often does not include patients who tested negative for SARS-CoV-2, which is important to control for health-seeking behaviors that may have changed during the pandemic as well as social factors that may have affected healthcare delivery and mental well-being for all individuals throughout the pandemic.[9]

Our primary objective was to measure the physical and mental quality of life outcomes in a cohort of patients who presented to EDs and were tested for SARS-CoV-2. The secondary objectives were to identify clinical, demographic, and sociocultural factors associated with the quality of life measures.

Methods

Study design and setting

This multicenter pan-Canadian study recruited patients from 22 EDs across five provinces that participated in the CCEDRRN (pronounced "sedrin") collaboration.[10–13] We recruited consecutive consenting patients between March 1, 2020, to July 14, 2021.[10] The University of British Columbia Research Ethics Board reviewed and approved the study protocol (H20-01015) with an exemption to obtain informed consent for retrospective registry enrollment, with permission to contact patients to obtain their consent for follow-up phone interviews. CCEDRRN's Patient Engagement Committee reviewed and provided significant input into the research question and edited consent and data collection forms to ensure readability and acceptability across diverse social, cultural, and ethnic contexts. They provided invaluable input in interpreting the study results, helping the authors contextualize results, and in writing the article.

Participants

We included consenting patients aged ≥17 years who presented to participating EDs and were tested for SARS-CoV-2.[14] We assigned patients to the exposure group if they had a laboratory-confirmed case of COVID-19, defined as one or more nucleic acid amplification tests positive for SARS-CoV-2 from specimens collected within the community in the 14 days before the ED visit, during the ED visit, or in the first five days after admission.[15] We assigned patients to the test-negative control group if all of their recorded SARS-CoV-2 tests were negative. We contacted consecutive SARS-CoV-2–positive patients by phone and matched consenting SARS-CoV-2–positive patient with up to four randomly selected SARS-CoV-2–negative patients (test-negative controls) who presented to the same site and within the same week. This allowed us to control for COVID-19 prevalence in the region, healthcare seeking behavior, as well as time trends in public health measures, which affected all patients and varied over the course of the pandemic. We recruited up to four test-negative controls per case to enable better matching on baseline variables.

Data collection

Trained research assistants abstracted data from paper-based and hospital electronic medical records. Research assistants conducted follow-up interviews contacting patients up to five attempts. We aimed to follow up patients at 30 days, 60 days, 6 months, and 12 months after their ED visit. Owing to varying delays in institutional ethics and privacy approvals, research assistants were only able to follow up patients who presented in the early pandemic (March 2020 – December 2020) at 200–395 days after the date of their ED or hospital discharge. These patients were generally only followed up once. Research assistants contacted all other patients for the first and second follow-up interviews per the approved protocol.

At follow-up, research assistants classified COVID-19 severity using the World Health Organization Ordinal Outcome Scale, a validated disease severity scale for COVID-19 patients assigned retrospectively for the patient's hospital visit.[16] Research assistants

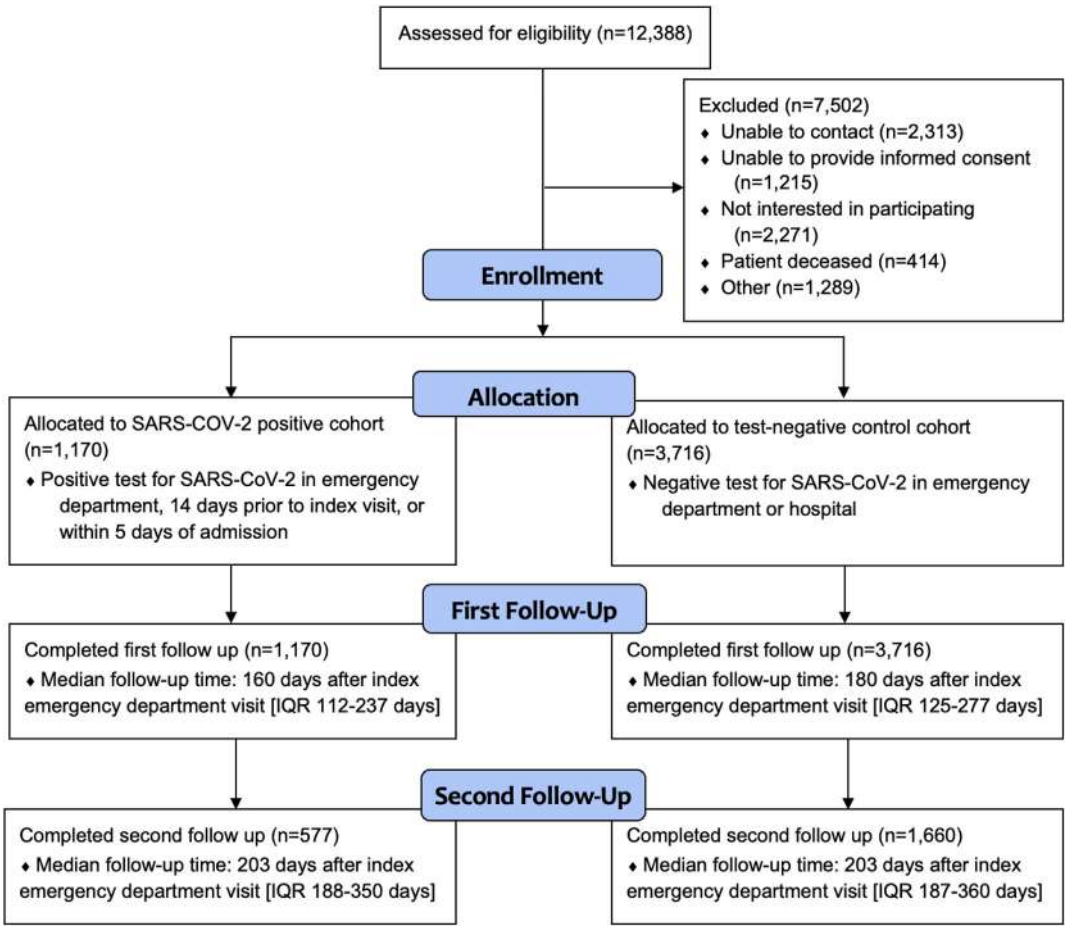


Fig. 1. Flow diagram of enrolled patients.

also ascertained sociocultural and demographic variables including sex, race, and education, COVID-19 vaccination status, as well as the quality of life using the Veterans RAND 12-item health survey, a health-related quality of life tool commonly used in American and Canadian patient-reported outcomes research.[17]

We converted the VR-12 survey responses to physical health component scores (PCS) and mental health component scores (MCS) based on patients' general health perception, physical functioning, role limitations, bodily pain, energy fatigue, social functioning, and perceived mental health.[17] PCS and MCS ranged between 0 and 100. The US population standards indicate mean scores of 50, with standard deviations (SDs) of 10.[17] Higher scores represent better self-reported health status, and an absolute difference of $\geq$1 point in either scores is considered socially and clinically relevant.[17] We collected data using REDCap (Vanderbilt University, Nashville, Tennessee, USA).

Outcome variables and covariates

Our two main outcome measures were PCS and MCS scores. We identified covariates from retrospective chart review and telephone follow-up and included demographic, clinical, and sociocultural variables (Supplemental Table 1).[10]

Statistical analysis

We summarized continuous covariates with median and interquartile range (IQR) statistics and categorical as percentages. We reported PCS and MCS using means and SD. We imputed missing data via multiple imputation using a fully conditional specification for five imputations.[18] We used propensity score matching to create a cohort of SARS-CoV-2–positive and SARS-CoV-2–negative patients matched on patient characteristics and clinical variables captured at the time of the ED visit.[19] We used logistic regression to model patients' propensity of testing positive for SARS-CoV-2 including the baseline covariates age, biological sex, province of residence, pandemic wave of presentation, respiratory distress on arrival, lowest oxygen saturation recorded in ED, oxygen requirements, comorbidities, ambulance arrival, in-hospital intubation, 7-day community incidence of COVID-19 in their area of residence, housing situation, tobacco and illicit substance use, race, immigrant status, employment status, education level, and income level. We matched SARS-CoV-2–positive patients with test-negative patients from the control group in a one-to-one ratio using a greedy neighbor approach without replacement and a caliper of 0.2 pooled SDs of the logit of the propensity score. The outcomes of patients that were not successfully matched were not analyzed.[20] We assessed residual differences in the baseline variables between groups using Student's *t*-test for continuous predictors and Chi-squared tests for categorical predictors. In the primary analyses, we used linear mixed effects models to model PCS and MCS adjusting for age, sex, all comorbidities, measures of oxygen requirements (presence of respiratory distress, oxygen requirements, and intubation), effect modification by age and immunization status, as well as the number of days between the index ED visit and follow-up interview.

Table 1
Baseline variables of patients.

Variable	SARS-CoV-2 positive ($n = 1170$)	SARS-CoV-2 negative ($n = 3716$)
Age, *n* (%)		
17–24 years	36 (3.1%)	213 (5.7%)
25–39 years	213 (18.2%)	771 (20.8%)
40–64 years	559 (47.8%)	1350 (36.3%)
65–79 years	301 (25.7%)	989 (26.6%)
>80 years	61 (5.2%)	393 (10.6%)
Sex, *n* (%)		
Male	581 (49.7%)	1752 (47.1%)
Female	589 (50.3%)	1964 (52.9%)
Arrival from, *n* (%)		
Home	1124 (96.1%)	3471 (93.4%)
Institutional living (long-term care/rehabilitation facility/interhospital transfer)	27 (2.3%)	168 (4.5%)
Homeless/correctional facility/other	19 (1.6%)	77 (2.1%)
Wave of presentation, *n* (%)		
Wave 1 (March 1, 2020 – June 30, 2020)	263 (22.5%)	682 (18.4%)
Wave 2 (July 1, 2020 – February 28, 2021)	460 (39.3%)	1619 (43.6%)
Wave 3 (March 1, 2021 – July 14, 2021)	447 (38.2%)	1415 (38.0%)
Province, *n* (%)		
Western Canada (BC and SK)	529 (45.2%)	1835 (49.4%)
Ontario	55 (4.7%)	197 (5.3%)
Eastern Canada (QC and NS)	586 (50.1%)	1684 (45.3%)
7-day average COVID-19 incident cases before the ED visit (%) per 100,000		
0–1.99	167 (14.3%)	628 (16.9%)
2–7.99	125 (10.7%)	528 (14.2%)
>8	878 (75.0%)	2560 (68.9%)
Lowest Oxygen Saturation in Emergency Department, median [IQR]	95 [5]	97 [3]
	Missing ($n = 9$)	Missing ($n = 59$)
Comorbid conditions, *n* (%)		
Secondary immunodeficiency (active malignancy, organ transplant recipient, severe liver disease)	50 (4.3%)	350 (9.4%)
Asthma	131 (11.2%)	341 (9.2%)
Atrial fibrillation	43 (3.7%)	271 (7.3%)
Chronic kidney disease	43 (3.7%)	205 (5.5%)
Chronic lung disease	56 (4.8%)	339 (9.1%)
Chronic neurological disorder	57 (4.9%)	363 (9.8%)
Congestive heart failure	29 (2.5%)	150 (4.0%)
Coronary artery disease	89 (7.7%)	392 (10.5%)
Diabetes	198 (16.9%)	543 (14.6%)
Dyslipidemia	249 (21.3%)	815 (21.9%)
Hypertension	379 (32.4%)	1229 (33.1%)
Hypothyroidism	71 (6.1%)	338 (9.1%)
Obesity (clinical impression)	36 (3.1%)	105 (2.8%)
Rheumatologic disorder	107 (9.1%)	471 (12.7%)
Past malignant neoplasm (cancer)	51 (4.4%)	272 (7.3%)
Psychiatric condition/Mental health diagnosis	119 (10.2%)	699 (18.9%)
Arrived by ambulance, *n* (%)		
Yes	488 (41.7%)	1208 (32.5%)
No	682 (58.3%)	2508 (67.5%)
Respiratory distress, *n* (%)		
Yes	247 (21.1%)	333 (9.0%)
No	923 (78.9%)	3383 (91.0%)
Tobacco use, *n* (%)		
Current	30 (2.5%)	391 (10.5%)
Past	77 (6.6%)	365 (9.8%)
Never	1063 (90.9%)	2960 (79.7%)
Illicit substance use, *n* (%)		
Yes	45 (3.8%)	348 (9.4%)
No	1125 (96.2%)	3368 (90.6%)
Oxygen required in ED, *n* (%)		
Yes	267 (22.8%)	352 (9.5%)
No	903 (77.2%)	3364 (90.5%)
Intubation in hospital, *n* (%)		
Yes	63 (5.4%)	97 (2.6%)
No	1107 (94.6%)	3619 (97.4%)
WHO ordinal scale assessment, *n* (%)		
Score 1	394 (33.7%)	1817 (48.9%)
Score 2	514 (43.9%)	1163 (31.3%)
Score 3	59 (5.0%)	402 (10.8%)
Score 4	117 (10.0%)	215 (5.8%)
Score 5	31 (2.6%)	17 (0.5%)
Score 6–7	44 (3.8%)	46 (1.2%)
Missing	11 (1.0%)	56 (1.5%)
Vaccination status, *n* (%)		
Not vaccinated	1131 (96.7%)	3445 (92.7%)

Table 1 (*continued*)

Variable	SARS-CoV-2 positive ($n = 1170$)	SARS-CoV-2 negative ($n = 3716$)
Partially/fully vaccinated	29 (2.5%)	253 (6.8%)
Missing	10 (0.8%)	18 (0.5%)
Race, *n* (%)		
Arab/Middle East	64 (5.5%)	158 (4.2%)
Black	45 (3.9%)	104 (2.8%)
East Asian/Southeast Asian	155 (13.2%)	315 (8.5%)
Indigenous	26 (2.2%)	105 (2.8%)
Latin American	49 (4.2%)	70 (1.9%)
South Asian	129 (11.0%)	186 (5.0%)
White	650 (55.6%)	2590 (69.7%)
Unknown/prefer not to answer	52 (4.4%)	188 (5.1%)
Current income bracket, *n* (%)		
<$22,440–29,900	168 (14.4%)	700 (18.8%)
$29,901–42,300	108 (9.2%)	377 (10.2%)
$42,301–55,300	72 (6.2%)	238 (6.4%)
$55,301–73,700	97 (8.3%)	295 (7.9%)
$73,701+	294 (25.1%)	839 (22.6%)
Prefer not to answer	431 (36.8%)	1267 (34.1%)
Immigrated to Canada, *n* (%)		
Yes	498 (42.6%)	1085 (29.2%)
No	627 (53.6%)	2521 (67.8%)
Prefer not to answer	45 (3.8%)	110 (3.0%)
Highest level of education achieved, *n* (%)		
No high school	85 (7.3%)	336 (9.0%)
High school	254 (21.7%)	879 (23.7%)
Trade certificate or diploma	92 (7.9%)	303 (8.2%)
College	148 (12.6%)	352 (9.5%)
University certificate or diploma	105 (9.0%)	387 (10.4%)
University degree	444 (37.9%)	1309 (35.2%)
Prefer not to answer	42 (3.6%)	150 (4.0%)
Employment, *n* (%)		
Employed	733 (62.7%)	1812 (48.8%)
Unemployed	114 (9.7%)	506 (13.6%)
Prefer not to Answer	39 (3.3%)	108 (2.9%)
Retired	284 (24.3%)	1290 (34.7%)

BC, British Columbia; ED, emergency department; IQR, interquartile range; NS, Nova Scotia; QC, Québec; SK, Saskatchewan.

We conducted two sensitivity analyses. The first modeled PCS and MCS jointly using multivariate analysis of variance for repeated outcomes to reflect the correlation between PCS and MCS scores from the same patient. The second included only patients who presented with recorded viral symptoms in the ED, including cough, shortness of breath, fever, chills, headache, nausea/vomiting, diarrhea, hemoptysis, chest pain, fatigue/malaise, myalgia, or dysgeusia/anosmia.[11]

As a secondary analysis, we examined demographic, clinical, and sociocultural risk factors for worse PCS and MCS using the entire sample of SARS-CoV-2 positive and test-negative patients.[21] We used principal component analysis to evaluate whether the large number of potentially collinear predictors could be reduced. We used multivariable linear regression to model the reduced number of predictors with PCS and MCS outcomes. We conducted a sensitivity analysis repeating the same modeling strategy for the SARS-CoV-2–positive patients only to identify differences in the type and strength of PCS and MCS risk factors between COVID-19 patients and the entire cohort.

We considered a *P*-value below 0.05 to be significant. We used Bonferroni's correction to adjust for multiple comparisons during analysis.[22] We analyzed data using R 4.1 (Vienna, Austria) and SAS 9.4 (SAS Institute Inc., Cary, USA).

Results

Of 12,388 patients enrolled in the CCEDRRN registry during the study period, 2739 were COVID-19–positive, and 9649 were test-negative controls. A total of 4886 patients (39%) consented to participate: SARS-CoV-2–positive patients comprised 24% of the sample ($n = 1170$), with 3716 test-negative controls (Fig. 1). A total of 2545 patients completed one follow-up interview, and 2341 completed two. The median follow-up time for the first interview was 169 days (IQR 112–237) for SARS-CoV-2–positive patients and 187 days (IQR 125–277) for test-negative controls (Fig. 1). The median follow-up time for the second interview was 203 days after the index ED visit for both groups with an IQR of 188–350 and 187–360 days for SARS-CoV-2–positive and test-negative groups, respectively.

For both SARS-CoV-2–positive and test-negative control groups, the distribution of baseline characteristics is presented in Table 1. The mean PCS was 42.4 (SD = 10.9) among SARS-CoV-2–positive patients and 40.7 (SD = 12.5) among test-negative controls. The mean MCS was 48.5 (SD = 11.2) among SARS-CoV-2–positive patients and 48.5 (SD = 11.5) among test-negative controls.

For the propensity-matched analysis, 1042 SARS-CoV-2–positive patients (89%) were matched with 1042 test-negative controls (28%). Matching was adequate for baseline characteristics and clinical measures (Table 2). The distribution of baseline characteristics for the unmatched sample is presented in Supplemental Table 2. The unadjusted mean PCS of matched patients was 42.4 (SD = 11.0) among SARS-CoV-2–positive patients and 41.8 (SD = 12.1) among test-negative controls. The unadjusted mean MCS was 48.3 (SD = 11.4) among SARS-CoV-2–positive patients and 49.5 (SD = 11.0) among test-negative controls (Fig. 2).

The effect of SARS-CoV-2 infection on PCS, adjusted for the time between the ED visit and interview, effect modification, and a priori confounders, was +0.50 (adjusted *P*-value = 0.51; 95% confidence interval [CI]: -0.36, 1.36). The adjusted effect of SARS-CoV-2 infection on MCS was -1.01 (adjusted *P*-value = 0.042; 95% CI: -1.91,

Table 2
Confounder variables that were matched using propensity score matching.

Variable	SARS-CoV-2 positive ($n = 1042$)	SARS-CoV-2 negative ($n = 1042$)	*P*-value
Age, *n* (%)			0.62
17–24 years	36 (3.5%)	40 (3.8%)	
25–39 years	199 (19.1%)	234 (22.5%)	
40–64 years	502 (48.2%)	445 (42.7%)	
65–79 years	249 (23.9%)	236 (22.6%)	
>80 years	56 (5.4%)	87 (8.3%)	
Sex, *n* (%)			0.90
Male	511 (49.0%)	515 (49.4%)	
Female	531 (51.0%)	527 (50.6%)	
Arrival from, *n* (%)			0.64
Home	999 (95.9%)	990 (95.0%)	
Institutional living (long-term care/rehabilitation facility/interhospital transfer)	26 (2.5%)	32 (3.1%)	
Homeless/correctional facility/other	17 (1.6%)	20 (1.9%)	
Lowest oxygen saturation in emergency department, median [IQR]	96 [4]	96 [4]	0.99
Wave of presentation, *n* (%)			0.82
Wave 1 (March 1, 2020 – June 30, 2020)	237 (22.7%)	249 (23.9%)	
Wave 2 (July 1, 2020 – February 28, 2021)	413 (39.6%)	405 (38.9%)	
Wave 3 (March 1, 2021 – July 14, 2021)	392 (37.6%)	388 (37.2%)	
Province, *n* (%)			0.98
Western Canada (BC and SK)	468 (44.9%)	464 (44.5%)	
Ontario	51 (4.9%)	52 (5.0%)	
Eastern Canada (QC and NS)	523 (50.2%)	526 (50.5%)	
7-day incidence, *n* (%), per 100,000			0.94
0–1.99	150 (14.4%)	204 (19.6%)	
2–7.99	116 (11.1%)	121 (11.6%)	
>8	776 (74.5%)	717 (68.8%)	
Comorbid conditions, *n* (%)			
Secondary immunodeficiency (active malignant neoplasm, organ transplant recipient, severe liver disease)	48 (4.6%)	50 (4.8%)	0.92
Asthma	109 (10.5%)	97 (9.3%)	0.42
Atrial fibrillation	40 (3.8%)	44 (4.2%)	0.74
Chronic kidney disease	41 (3.9%)	42 (4.0%)	1.00
Chronic lung disease	53 (5.1%)	60 (5.8%)	0.56
Chronic neurological disorder	55 (5.3%)	61 (5.9%)	0.63
Congestive heart failure	27 (2.6%)	31 (3.0%)	0.69
Coronary artery disease	82 (7.9%)	90 (8.6%)	0.58
Diabetes	172 (16.5%)	169 (16.2%)	0.91
Dyslipidemia	216 (20.7%)	215 (20.6%)	0.65
Hypertension	324 (31.1%)	314 (30.1%)	0.67
Hypothyroidism	69 (6.6%)	63 (6.0%)	1.00
Obesity (clinical impression)	31 (3.0%)	31 (3.0%)	1.00
Rheumatologic disorder	92 (8.8%)	97 (9.3%)	0.76
Past malignant neoplasm (cancer)	48 (4.6%)	45 (4.3%)	0.83
Psychiatric condition/mental health diagnosis	110 (10.6%)	123 (11.8%)	0.40
Arrived by ambulance, *n* (%)			0.79
Yes	403 (38.7%)	396 (38.0%)	
No	639 (61.3%)	646 (62.0%)	
Respiratory distress, *n* (%)			0.95
Yes	170 (16.3%)	168 (16.1%)	
No	872 (83.7%)	874 (83.9%)	
Tobacco use, *n* (%)			0.98
Current	30 (2.9%)	31 (3.0%)	
Past	71 (6.8%)	73 (7.0%)	
Never	941 (90.3%)	938 (90.0%)	
Illicit substance use, *n* (%)			0.22
Yes	44 (4.2%)	57 (5.5%)	
No	998 (95.8%)	985 (94.5%)	
Oxygen required in ED, *n* (%)			0.91
Yes	170 (16.3%)	167 (16.0%)	
No	872 (83.7%)	875 (84.0%)	
Intubation in hospital, *n* (%)			1.00
Yes	40 (3.8%)	39 (3.7%)	
No	1002 (96.2%)	1003 (96.3%)	
Race, *n* (%)			0.91
Arab/Middle East	66 (6.3%)	70 (6.7%)	
Black	45 (4.3%)	50 (4.8%)	
East Asian/Southeast Asian	125 (12.0%)	108 (10.4%)	
Indigenous	24 (2.3%)	26 (2.5%)	
Latin American	45 (4.3%)	40 (3.8%)	
South Asian	112 (10.7%)	114 (10.9%)	
White	625 (60.0%)	634 (60.8%)	
Current income bracket, *n* (%)			0.74
<$22,440–29,900	253 (24.3%)	234 (22.5%)	
$29,901–42,300	162 (15.5%)	153 (14.7%)	

Table 2 (*continued*)

Variable	SARS-CoV-2 positive (n = 1042)	SARS-CoV-2 negative (n = 1042)	P-value
$42,301–55,300	102 (9.8%)	104 (10.0%)	
$55,301–73,700	130 (12.5%)	145 (13.9%)	
$73,701+	395 (37.9%)	406 (39.0%)	
Immigrated to Canada, *n* (%)			0.93
Yes	439 (42.1%)	436 (41.8%)	
No	603 (57.9%)	606 (58.2%)	
Highest level of education achieved, *n* (%)			0.96
No high school	79 (7.6%)	74 (7.1%)	
High school	236 (22.6%)	223 (21.4%)	
Trade certificate or diploma	86 (8.3%)	83 (8.0%)	
College	127 (12.2%)	137 (13.1%)	
University certificate or diploma	97 (9.3%)	99 (9.5%)	
University degree	417 (40.0%)	426 (40.9%)	
Employment, *n* (%)			0.97
Employed	670 (64.3%)	675 (64.8%)	
Unemployed	111 (10.7%)	110 (10.6%)	
Retired	261 (25.0%)	257 (24.7%)	

BC, British Columbia; ED, emergency department; IQR, interquartile range; NS, Nova Scotia; QC, Québec; SK, Saskatchewan.

-0.11). We reported similar findings for the adjusted joint analysis of PCS and MCS (Supplemental Table 3).

We completed a sensitivity analysis with patients who presented with documented SARS-CoV-2 symptoms (n = 936). The significant difference in MCS was maintained through our sensitivity analysis (Table 3).

We summarized the risk factors for PCS and MCS in Table 4 for all 4886 patients enrolled in this study. The principal component analyses identified two principal components representing oxygen requirements (comprised lowest oxygen saturation in the ED, presence of respiratory distress, intubation, and supplemental oxygen requirements) and medically complex comorbidities (comprised secondary immunodeficiency, chronic neurological disorder, rheumatological disease, and history of past malignancy). The two principal components were analyzed as predictors for PCS and MCS in Table 4. The results of the sensitivity analysis for SARS-CoV-2–positive patients only are summarized in Supplemental Table 4.

Discussion

Using a propensity score–matched design, we found that COVID-19 patients who presented to EDs reported lower mental health status during the follow-up period. Interestingly, the physical health outcomes of COVID-19 patients were similar to those of other patients who presented to the ED at the same time during the pandemic.

Although studies that examined physical health outcomes of COVID-19 survivors to known population norms reported significantly worse outcomes, those studies did not control for important confounders such as the occurrence of complex comorbidities, age distributions, or health services access variables between

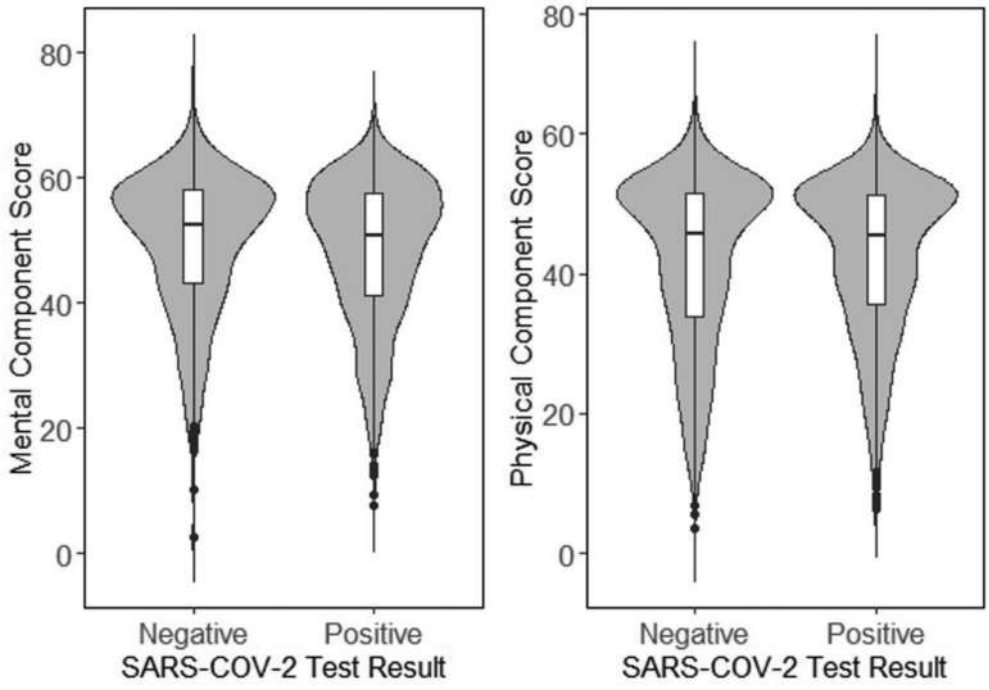


Fig. 2. Violin plots of physical and mental component scores across COVID-19 cases and test-negative controls after propensity score matching.

Table 3
Subgroup analysis excluding 936 patients without respiratory-related symptoms.

Variable	*Physical component score*		*Mental component score*	
	Beta	*P*-value	Beta	*P*-value
COVID-19 positive	0.30	0.49	−1.07	0.01
Time between emergency department visit and interview (days)	0.002	0.33	0.002	0.26
Age (years)	−0.12	<0.001	0.08	<0.001
Female sex	−2.09	<0.001	−2.62	<0.001
Smoking				
Never	2.13	0.10	1.56	0.24
Past	2.51	0.10	1.66	0.28
Intubated	−0.96	0.44	−1.30	0.30
Respiratory distress	−0.59	0.35	0.14	0.83
Supplemental oxygen required in emergency department	−1.41	0.03	−0.29	0.67
Comorbidities				
Asthma	−1.62	0.02	0.08	0.91
Atrial fibrillation	−4.39	<0.001	0.30	0.80
Chronic kidney disease	−2.87	0.02	−0.29	0.81
Chronic lung disease	−3.79	<0.001	−1.45	0.15
Chronic neurological disorder	−0.73	0.45	−0.68	0.48
Congestive heart failure	−4.12	0.004	−1.41	0.33
Coronary artery disease	−1.43	0.11	−1.30	0.15
Diabetes	−2.34	<0.001	−1.19	0.07
Dyslipidemia	−0.26	0.68	−0.12	0.85
Hypertension	0.61	0.31	0.32	0.60
Hypothyroidism	−0.62	0.49	0.72	0.43
Obesity	−0.77	0.57	1.04	0.45
Past malignant neoplasm	0.09	0.92	2.82	0.006
Psychiatric condition/mental health diagnosis	−1.26	0.08	−6.42	<0.001
Rheumatological disorder	−3.57	<0.001	−0.46	0.55
Secondary immunodeficiency	−3.00	0.004	1.79	0.09
Immunized to SARS-CoV-2	4.23	0.31	−4.47	0.29
Immunized to SARS-CoV-2 × age (interaction)	−0.08	0.21	−0.05	0.46

populations.[23,24] However, studies examining the physical health of COVID-19 patients that controlled for such factors reported similar outcomes as this study; Huang et al. observed no significant difference in the physical quality of life of hospitalized COVID-19 survivors with severe disease compared with ambulatory COVID-19 survivors with mild disease.[25] Our findings may be explained by considering access barriers to care that occurred in the early waves of the pandemic. Non-COVID patients were reluctant to come to EDs due to fear of contracting the virus and may have suffered from complications and challenges related to the delay in seeking care, impacting self-reported physical health.[26] Our results indicate that the physical health outcomes reported by COVID-19 survivors are comparable to those experienced by non-COVID patients presenting to EDs with acute health-related challenges at similar time points during the pandemic.

This study found significantly lower MCS among COVID-19 survivors compared with matched test-negative controls and is consistent with case−control studies investigating the direct effects of COVID-19 on patients who reported suffering from post-traumatic stress, anxiety, and depression postinfection.[27,28] Because of the poor self-reported mental health outcomes observed among COVID-19 patients in this study, there is a need to investigate and understand the specific factors associated with worse mental health outcomes for this population. This information may be useful to inform targeted mental health services among COVID-19 survivors; recommending outpatient mental health services may be one strategy to address the suboptimal mental health status of post-COVID patients. However, in the Canadian context, this is complicated by the fact that many subacute mental health care services are not publicly funded and not universally available.

Several clinical, demographic, and sociocultural characteristics influenced patient's self-reported physical and mental quality of life. Not surprisingly, factors such as severe disease, oxygen requirements, smoking, and age were associated with worse physical health outcomes. Factors such as lower annual income, having no fixed address, illicit substance use, and having a history of a psychiatric or mental health diagnosis, were associated with worse mental health outcomes. Interventions and policies that aim to support COVID-19 survivors' health should reflect the differential impacts of clinical, demographic, and sociocultural factors on physical and mental health outcomes.

This study was not without limitations. We only interviewed patients by phone because of restrictions with in-person research at hospital sites. This necessarily excluded individuals without phone access and those unable to use phones. This may have led us to underestimate the impact of COVID-19 on the health of disadvantaged or vulnerable populations. Furthermore, we focused solely on two composite outcome measures of physical and mental health. Single numerical scores are effective for identifying trends in aggregate but cannot provide details about different dimensions of physical and mental health at an individual level.[29] This study was initiated before the World Health Organization defined the post-COVID condition, which prevented us from determining whether patients meeting the diagnostic criteria for the post-COVID condition had worse physical or mental health outcomes compared with patients who did not meet the diagnostic criteria for the post-COVID condition. Finally, our results are only representative of consenting patients who sought care in an ED. Because a large proportion of COVID-19 patients did not seek evaluation or medical treatment in EDs, we cannot assume these findings apply more generally to patients with COVID-19 who did not access ED care.

The strengths of this study include incorporation of data from the largest consecutive cohort of COVID-19 patients presenting to academic, non-academic, urban, and remote EDs, which were included in the CCEDRRN registry. Given that most of the available literature on COVID-19 outcomes are from studies that neither control for social or demographic characteristics of patients, nor

Table 4
Principal component regression analysis for SARS-CoV-2–positive and test-negative control patients.

Principal component (PC)/variable	Mental component score			Physical component score		
	Beta	95% confidence interval	*P*-value	Beta	95% confidence interval	*P*-value
Oxygen requirements (PC1)	0.08	−0.37, 0.21	0.58	−0.31	−0.59, −0.03	0.03
Medically complex comorbidities (PC2)	0.25	−0.57, 0.07	0.13	−1.5	−1.8, −1.2	<0.001
Time between emergency department visit and interview (days)	0	−0.01, 0.00	0.9	0.01	0.00, 0.01	0.037
SARS-CoV-2						
Negative	–	–		–	–	
Positive	−0.11	−0.91, 0.68	0.8	1.9	1.2, 2.7	<0.001
Smoking						
Never	–	–		–	–	
Current	−2.5	−3.7, −1.3	<0.001	−1.6	−2.7, −0.40	0.008
Past	0.47	−0.65, 1.6	0.4	0.5	−0.58, 1.6	0.4
Illicit substance use						
Never	–	–		–	–	
Current	−2.4	−3.6, -1.1	<0.001	−0.18	−1.4, 1.0	0.8
Age (years)	0.07	−0.01, 0.14	0.08	−0.15	−0.23, −0.08	<0.001
Sex						
Male	–	–		–	–	
Female	−2.3	−2.9, −1.6	<0.001	−1.6	−2.2, −1.0	<0.001
Race						
White	–	–		–	–	
Arab/Middle East	−1.3	−2.9, 0.29	0.11	−1.2	−2.7, 0.39	0.14
Black	1	−0.88, 2.9	0.3	0.29	−1.5, 2.1	0.8
East Asian/Southeast Asian	−0.34	−1.6, 0.89	0.6	0.34	−0.85, 1.5	0.6
Indigenous	−0.1	−2.0, 1.8	>0.9	−0.73	−2.6, 1.1	0.4
Latin American	−0.68	−2.8, 1.4	0.5	−1	−3.0, 1.0	0.3
South Asian	1	−0.44, 2.4	0.2	0.18	−1.2, 1.5	0.8
Income						
<$22,440	−1.6	−2.7, -0.59	0.002	−1.6	−2.7, −0.64	0.001
$22,401–29,900	−1.3	−2.5, 0.00	0.05	−1.4	−2.6, −0.17	0.025
$29,901–36,200	−0.26	−1.5, 1.0	0.7	−1	−2.2, 0.24	0.11
$36,201–42,300	−0.76	−2.1, 0.56	0.3	−0.66	−1.9, 0.61	0.3
$42,301–48,400	−1.8	−3.4, −0.22	0.026	−1.2	−2.8, 0.33	0.12
$48,401–55,300	−0.78	−2.3, 0.73	0.3	−1.6	−3.0, −0.10	0.036
$55,301–63,200	−0.07	−1.5, 1.4	>0.9	−0.83	−2.2, 0.55	0.2
$63,201–73,700	1.3	−0.08, 2.7	0.065	−0.9	−2.2, 0.44	0.2
$73,701–91,100	0.71	−0.47, 1.9	0.2	0.06	−1.1, 1.2	>0.9
$91,100+	–	–		–	–	
Immigrated to Canada						
No	–	–		–	–	
Yes	0.17	−0.72, 1.1	0.7	1.3	0.47, 2.2	0.002
Education						
No high school	0.39	−1.0, 1.8	0.6	0.31	−1.1, 1.7	0.7
High school	0.85	−0.28, 2.0	0.14	0.64	−0.45, 1.7	0.2
Trade certificate or diploma	−0.19	−1.6, 1.2	0.8	0.04	−1.3, 1.4	>0.9
College	0.14	−1.2, 1.5	0.8	1.4	0.14, 2.7	0.03
University certificate or diploma	0.68	−0.40, 1.8	0.2	1.6	0.55, 2.6	0.003
University degree	–	–		–	–	
Wave of presentation						
Wave 1	–	–		–	–	
Wave 2	0.18	−1.0, 1.4	0.8	2.6	1.5, 3.8	<0.001
Wave 3	−0.25	−1.6, 1.1	0.7	2.2	0.88, 3.5	<0.001
Province						
Eastern Canada	–	–		–	–	
Ontario	−0.01	−1.6, 1.6	>0.9	−0.51	−2.0, 1.0	0.5
Western Canada	−0.29	−1.2, 0.57	0.5	0.87	0.04, 1.7	0.039
7-day community incidence per 100,000	0.03	0.00, 0.06	0.05	−0.01	−0.04, 0.02	0.5
Arrival from						
Home	–	–		–	–	
Institutional living	1.2	−0.46, 2.8	0.2	−0.72	−2.3, 0.85	0.4
Homeless/correctional facility/other	−2.3	−4.6, 0.06	0.056	−0.62	−2.8, 1.6	0.6
Arrived by ambulance						
No	–	–		–	–	
Yes	0.17	−0.55, 0.88	0.6	−1.8	−2.5, −1.1	<0.001
Comorbidities						
Asthma	−0.78	−1.8, 0.27	0.14	−1.7	−2.7, −0.65	0.001
Atrial fibrillation	0.08	−1.3, 1.4	>0.9	−0.92	−2.2, 0.39	0.2
Chronic kidney disease	0.23	−1.3, 1.7	0.8	−1.8	−3.2, −0.34	0.015
Chronic lung disease	−0.22	−1.5, 1.0	0.7	−3.8	−5.0, −2.6	<0.001
Congestive heart failure	0.48	−1.3, 2.2	0.6	−4.5	−6.2, −2.8	<0.001
Coronary artery disease	−0.69	−1.8, 0.47	0.2	−0.78	−1.9, 0.32	0.2
Diabetes	−1	−1.9, −0.01	0.048	−1.1	−2.0, −0.13	0.025
Dyslipidemia	0.33	−0.58, 1.2	0.5	−0.01	−0.88, 0.86	>0.9

(*continued on next page*)

Table 4 (*continued*)

Principal component (PC)/variable	Mental component score			Physical component score		
	Beta	95% confidence interval	*P*-value	Beta	95% confidence interval	*P*-value
Hypertension	0.02	−0.82, 0.86	>0.9	−0.19	−1.0, 0.61	0.6
Hypothyroidism	−0.32	−1.5, 0.83	0.6	−0.53	−1.6, 0.58	0.3
Obesity	0.69	−1.2, 2.6	0.5	−1.9	−3.7, −0.10	0.039
Psychiatric condition/mental health diagnosis	−5.4	−6.3, −4.5	<0.001	−0.09	−0.93, 0.75	0.8
Immunized to SARS-CoV-2						
Yes	–	–		–	–	
No	0.2	−4.7, 5.1	>0.9	−2.1	−6.8, 2.6	0.4
Immunized to SARS-CoV-2 × age (*interaction*)	0.01	−0.07, 0.08	0.8	0.03	−0.04, 0.10	0.4
World Health Organization ordinal outcome score						
Score 1	–	–		–	–	
Score 2	−2.8	−3.5, −2.0	<0.001	−6.7	−7.4, −6.0	<0.001
Score 3	−1.4	−2.5, −0.21	0.021	−4.4	−5.5, −3.3	<0.001
Score 4	−1.8	−3.2, −0.41	0.011	−5.5	−6.8, −4.1	<0.001
Score 5	−2.6	−5.8, 0.61	0.11	−3.3	−6.4, −0.20	0.037
Score 6–7	−0.08	−2.5, 2.4	>0.9	−7.4	−9.8, −5.1	<0.001

include test-negative comparison groups, this study strengthens the knowledge about patient-reported outcomes among COVID-19 survivors by reporting results that control for health-seeking behaviors and time trends in addition to clinical, demographic, and sociocultural variables.

With increasing case counts of COVID-19 throughout the world, it is essential for policy makers to develop comprehensive strategies to mitigate the long-term outcomes of COVID-19 patients. These findings may be useful for identifying individuals at greater risk for worse mental health outcomes post-COVID-19 and may guide the development of interventions and policies to better support COVID-19 survivors.

Author statements

Acknowledgments

The authors thank the University of British Columbia clinical coordinating center, legal, ethics, privacy, and contract staff. The authors thank the Kazis Lab of Boston University for providing access to the Veterans RAND 12-item health survey and scoring algorithm. The authors would like to thank the CCEDRRN patient engagement committee, whose perspectives, opinions, and feedback helped them critically interpret the results of this thesis. Finally, the authors want to acknowledge their colleagues in medicine, nursing, and the allied health professions who have been on the front lines of this pandemic from day 1 staffing the ambulances, emergency departments, intensive care units, and hospitals, bravely facing the risks of COVID-19 to look after their fellow citizens and after one another.

Fundings

The Canadian Institutes of Health Research (447,679, 464,947, and 466,880), Ontario Ministry of Colleges and Universities (C-655-2129), Saskatchewan Health Research Foundation (5357), Genome BC (COV024 and VAC007), Fondation du CHU de Québec (Octroi No. 4007), and Sero-Surveillance and Research (COVID-19 Immunity Task Force Initiative) provided peer-reviewed funding. The BC Academic Health Science Network and BioTalent Canada provided non–peer-reviewed funding. These organizations are not-for-profit and had no role in study conduct, analysis, or manuscript preparation.

Competing interests

None.

Author contributions

This article was derived from R.B.'s master's thesis in population and public health supervised by C.H., J.S., and R.M. R.B., C.H., and J.S. conceptualized the study design. R.B. analyzed the data and wrote the original draft. The writing proposal was reviewed by the CCEDRRN Publication and Protocol Review Committee and Data Access and Management Committee before release of data. P.A., C.H., L.G., M.L., and J.H. contributed to data collection at participating CCEDRRN sites. T.R. and L.G. provided patient-partner input in the interpretation of the study results. All authors contributed to data interpretation and to the review and editing of the article. All authors had final responsibility for the decision to submit for publication.

Data sharing

For investigators who wish to access data from the Canadian COVID-19 Emergency Department Rapid Response Network, proposals may be submitted to the network for review and approval by the network's peer-review publication committee, the data access and management committee, and the executive committee, as per the network's governance. Information regarding submitting proposals and accessing data may be found at https://ccedrrn.com/.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2022.11.016.

References

1. Wu X, Liu X, Zhou Y, Yu H, Li R, Zhan Q, et al. 3-month, 6-month, 9-month, and 12-month respiratory outcomes in patients following COVID-19-related hospitalisation: a prospective study. *Lancet Respir Med* 2021;**9**(7):747–54.
2. Salerno S, Sun Y, Morris EL, He X, Li Y, Pan Z, et al. Comprehensive evaluation of COVID-19 patient short- and long-term outcomes: disparities in healthcare utilization and post-hospitalization outcomes. *PLoS One* 2021;**16**(10):e0258278.
3. Boscolo-Rizzo P, Guida F, Polesel J, Marcuzzo AV, Capriotti V, D'Alessandro A, et al. Sequelae in adults at 12 months after mild-to-moderate coronavirus disease 2019 (COVID-19). *Int Forum Allergy Rhinol* 2021;**11**(12):1685–8.
4. Xiong Q, Xu M, Li J, Liu Y, Zhang J, Xu Y, et al. Clinical sequelae of COVID-19 survivors in Wuhan, China: a single-centre longitudinal study. *Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis* 2021;**27**(1):89–95.
5. Taquet M, Geddes JR, Husain M, Luciano S, Harrison PJ. 6-month neurological and psychiatric outcomes in 236 379 survivors of COVID-19: a retrospective cohort study using electronic health records. *Lancet Psychiatr* 2021;**8**(5):416–27.

6. Wynberg E, van Willigen HDG, Dijkstra M, Boyd A, Kootstra NA, van den Aardweg JG, et al. Evolution of COVID-19 symptoms during the first 12 months after illness onset. *Clin Infect Dis Off Publ Infect Dis Soc Am* 2021:ciab759.
7. Shoucri SM, Purpura L, DeLaurentis C, Adan MA, Theodore DA, Irace AL, et al. Characterising the long-term clinical outcomes of 1190 hospitalised patients with COVID-19 in New York City: a retrospective case series. *BMJ Open* 2021;**11**(6):e049488.
8. Haraldstad K, Wahl A, Andenæs R, Andersen JR, Andersen MH, Beisland E, et al. A systematic review of quality of life research in medicine and health sciences. *Qual Life Res Int J Qual Life Asp Treat Care Rehabil* 2019;**28**(10):2641–50.
9. Vandenbroucke JP, Brickley EB, Vandenbroucke-Grauls CMJE, Pearce N. A test-negative design with additional population controls can Be used to rapidly study causes of the SARS-CoV-2 epidemic. *Epidemiol Camb Mass* 2020;**31**(6):836–43.
10. Hohl CM, Rosychuk RJ, McRae AD, Brooks SC, Archambault P, Fok PT, et al. Development of the Canadian COVID-19 emergency department rapid response network population-based registry: a methodology study. *CMAJ Open* 2021;**9**(1):E261–70.
11. McRae AD, Hohl CM, Rosychuk R, Vatanpour S, Ghaderi G, Archambault PM, et al. CCEDRRN COVID-19 Infection Score (CCIS): development and validation in a Canadian cohort of a clinical risk score to predict SARS-CoV-2 infection in patients presenting to the emergency department with suspected COVID-19. *BMJ Open* 2021;**11**(12):e055832.
12. Hohl CM, Rosychuk RJ, Hau JP, Hayward J, Landes M, Yan JW, et al. Treatments, resource utilization, and outcomes of COVID-19 patients presenting to the emergency department across pandemic waves. Can J Emerg Med [Internet]. 2022 Apr 1 [cited 2021 Nov 23]; Available from: http://medrxiv.org/lookup/doi/10.1101/2021.07.30.21261288.
13. Hohl CM, Rosychuk RJ, Archambault PM, O'Sullivan F, Leeies M, Mercier É, et al. The CCEDRRN COVID-19 Mortality Score to predict death among nonpalliative patients with COVID-19 presenting to emergency departments: a derivation and validation study. *CMAJ Open* 2022;**10**(1):E90–9.
14. Jin YH, Cai L, Cheng ZS, Cheng H, Deng T, Fan YP, et al. A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version). *Mil Med Res* 2020;**7**(1):4.
15. Elimian KO, Ochu CL, Ebhodaghe B, Myles P, Crawford EE, Igumbor E, et al. Patient characteristics associated with COVID-19 positivity and fatality in Nigeria: retrospective cohort study. *BMJ Open* 2020;**10**(12):e044079.
16. de Terwangne C, Laouni J, Jouffe L, Lechien JR, Bouillon V, Place S, et al. Predictive accuracy of COVID-19 world health organization (WHO) severity classification and comparison with a bayesian-method-based severity score (EPI-SCORE). *Pathog Basel Switz* 2020;**9**(11):E880.
17. Snedden TR, Scerpella J, Kliethermes SA, Norman RS, Blyholder L, Sanfilippo J, et al. Sport and physical activity level impacts health-related quality of life among collegiate students. *Am J Health Promot AJHP* 2019;**33**(5):675–82.
18. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;**338**:b2393.
19. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011;**46**(3):399–424.
20. Rassen JA, Shelat AA, Myers J, Glynn RJ, Rothman KJ, Schneeweiss S. One-to-many propensity score matching in cohort studies. *Pharmacoepidemiol Drug Saf* 2012;**21**(Suppl 2):69–80.
21. Huang T, Li J, Zhang W. Application of principal component analysis and logistic regression model in lupus nephritis patients with clinical hypothyroidism. *BMC Med Res Methodol* 2020;**20**(1):99.
22. Chen SY, Feng Z, Yi X. A general introduction to adjustment for multiple comparisons. *J Thorac Dis* 2017;**9**(6):1725–9.
23. Qu G, Zhen Q, Wang W, Fan S, Wu Q, Zhang C, et al. Health-related quality of life of COVID-19 patients after discharge: a multicenter follow-up study. *J Clin Nurs* 2021;**30**(11–12):1742–50.
24. Bardakci MI, Ozturk EN, Ozkarafakili MA, Ozkurt H, Yanc U, Yildiz Sevgi D. Evaluation of long-term radiological findings, pulmonary functions, and health-related quality of life in survivors of severe COVID-19. *J Med Virol* 2021;**93**(9):5574–81.
25. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet Lond Engl* 2020;**395**(10223):497–506.
26. Lazzerini M, Barbi E, Apicella A, Marchetti F, Cardinale F, Trobia G. Delayed access or provision of care in Italy resulting from fear of COVID-19. *Lancet Child Adolesc Health* 2020;**4**(5):e10–1.
27. Kyzar EJ, Purpura LJ, Shah J, Cantos A, Nordvig AS, Yin MT. Anxiety, depression, insomnia, and trauma-related symptoms following COVID-19 infection at long-term follow-up. *Brain Behav Immun - Health* 2021;**16**:100315.
28. Ismael F, Bizario JCS, Battagin T, Zaramella B, Leal FE, Torales J, et al. Post-infection depressive, anxiety and post-traumatic stress symptoms: a prospective cohort study in patients with mild COVID-19. *Prog Neuro-Psychopharmacol Biol Psychiatry* 2021;**111**:110341.
29. Kronzer VL, Jerry MR, Ben Abdallah A, Wildes TS, McKinnon SL, Sharma A, et al. Changes in quality of life after elective surgery: an observational study comparing two measures. *Qual Life Res Int J Qual Life Asp Treat Care Rehabil* 2017;**26**(8):2093–102.

Public Health 215 (2023) 31–38

Contents lists available at ScienceDirect

Public Health

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

Original Research

Prevalence and epidemiological trends in mortality due to COVID-19 in Saudi Arabia

D.A. Alissa [a, b], W. Aburas [c], R. Almasuood [a], H.Y. Almudaiheem [a], Z. Al Aseri [d, e], F. Alrabiah [f], H. Ezzat [g, h, i], A.A. Moulana [j], M.M. Alawi [k], E. Al-mutairy [f], T. Alaama [a, b], M.S. Alamri [a], M.S. Bamousa [a, l], A.A. Alshehri [m], M.H. Alosaimi [a, n], A. Alali [o], R. Nori [p], E.B. Devol [q], G. Mohamed [q], A.H. Al-jedai [a, b, *]

[a] Therapeutic Affairs Deputyship, Ministry of Health, Riyadh, Saudi Arabia
[b] Colleges of Medicine and Pharmacy, Al-Faisal University, Riyadh, Saudi Arabia
[c] University of Hail, College of Pharmacy, Hail, Saudi Arabia
[d] Department of Emergency Medicine and Critical Care, King Saud University, Riyadh, Saudi Arabia
[e] Adult Critical Care, Ministry of Health, Riyadh, Saudi Arabia
[f] Department of Medicine, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia
[g] Hematology Programs, Ministry of Health, Saudi Arabia
[h] John Hopkins Aramco Healthcare, Dharan, Saudi Arabia
[i] University of British Columbia, Canada
[j] Forensic Medicine Center, Makkah Al Mukarramah, Saudi Arabia
[k] Infection Control & Environmental Health Unit, Department of Medical Microbiology & Parasitology, Faculty of Medicine, King Abdulaziz University Hospital, Jeddah, Saudi Arabia
[l] Forensic Medicine General Directorate, Ministry of Health, Riyadh, Saudi Arabia
[m] Forensic Medicine Center, Ministry of Health, Riyadh, Saudi Arabia
[n] Mortality Affairs, Ministry of Health, Riyadh, Saudi Arabia
[o] Internal Medicine and Infectious Diseases, King Saud Medical City, Riyadh, Saudi Arabia
[p] Internal Medicine Department, Prince Mohammed Bin Abdulaziz Hospital, Riyadh, Saudi Arabia
[q] Department of Biostatistics, Epidemiology & Scientific Computing, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia

ARTICLE INFO

Article history:
Received 22 May 2022
Received in revised form
17 July 2022
Accepted 19 July 2022
Available online 10 January 2023

Keywords:
COVID-19
Case fatality rate
Mortality
Prevalence
Epidemiological trends
Kingdom of Saudi Arabia

ABSTRACT

Objectives: This article describes the prevalence and epidemiological trends of COVID-19 mortality in the largest registry in the Kingdom of Saudi Arabia (KSA).
Study design: A prospective epidemiological cohort study using data from all healthcare facilities in KSA collected between March 23, 2020, and April 30, 2022. Data on the number of daily deaths directly related to COVID-19 were gathered, analyzed, and reported.
Method: Data analysis was carried out using national and regional crude case fatality rate and death per 100,000 population. Descriptive statistics using numbers and proportions were used to describe age, gender, nationality, and comorbidities. The mortality trend was plotted and compared with international figures. In addition, the most common comorbidities associated with mortality and the proportion of patients who received COVID-19 vaccine were reported.
Results: The total reported number of deaths between March 23, 2020, and April 30, 2022, was 9085. Crude case fatality rate was 1.21%, and death per 100,000 population was 25.38, which compared favorably to figures reported by several developed countries. The highest percentages of deaths were among individuals aged between 60 and 69 years, males (71%), and individuals with diabetes (60%). Only 2.8% of mortalities occur in patients who received COVID-19 vaccine. Diabetes, hypertension, and heart failure had the highest attributable risk of mortality among patients who died due to COVID-19.
Conclusion: Case fatality rate and death per 100,000 population in KSA are among the lowest in the world due to multiple factors. Several comorbidities have been identified, namely, diabetes, hypertension, obesity, and cardiac arrhythmias.

* Corresponding author. Ministry of Health, Riyadh, Saudi Arabia.
E-mail address: ahaljedai@moh.gov.sa (A.H. Al-jedai).

https://doi.org/10.1016/j.puhe.2022.07.014
0033-3506/

Introduction

COVID-19, an infectious respiratory illness caused by acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has devastating implications on health, economy, and daily life. The first case of COVID-19 emerged in Wuhan, China, in late December of 2019. Cases continued to mount worldwide, with the first confirmed case announced in the Kingdom of Saudi Arabia (KSA) in early March 2020.[1] COVID-19 was declared a pandemic by the World Health Organization (WHO) in March 2020. By mid-April 2022, approximately 500 million COVID-19–confirmed cases were reported worldwide, with 6,190,349 related deaths.[2,3]

In 2018, 2019, 2020, and 2021 death rates reported in KSA were 3.470, 3.513, 3.557, and 3.600, respectively. Circulatory system–related deaths were the highest contributor with 14.2%, followed by 4.3% for respiratory system diseases, 3.3% for certain infectious and parasitic diseases, and 2.2% for endocrine (nutritional and metabolic disorders).[4] The annual seasonal influenza epidemic contributed to death globally, accounting for approximately 290,000 to 650,000 respiratory deaths globally. High-risk populations, such as elderly patients and individuals with immunosuppressive conditions, chronic cardiac, pulmonary, and metabolic diseases, are more prone to death due to seasonal influenza.[5] Influenza and COVID-19 tend to overlap histologically, and both exhibit similar responses to severe virus-induced lung injury.[6]

To properly evaluate the impact of COVID-19 on society, excess deaths have been recognized as an essential and meaningful measure. It was reported as the third leading cause of death in the United States, right after heart diseases and cancer.[7] It has been suggested that the mortality rate because of COVID-19 could be underestimated or overestimated depending on reporting mechanism and definition of COVID-19 direct cause of mortality.[8] Healthcare systems were unprepared for the pandemic, where some countries suffered from overcrowded hospitals.[9,10] As a result, undercounting COVID-19 cases and their related deaths on a global level has been suggested by several investigators. This discrepancy may have been due to various factors, including but not limited to low laboratory capacity and inaccurate sampling techniques.[11,12] On the other hand, overestimation could result from counting deaths as COVID-19 related when in fact they are not. This could be because of testing patients who do not have the typical phenotype of viral illness caused by SARS-CoV-2 and who died of a different immediate cause of death, for example, motor vehicle accident or advanced cancer.[13] Disease severity may be attributed to multiple factors, pre-existing comorbidities, age, air pollution, temperature, and humidity.[14–17]

Excess mortality estimates the degree to which currently measured mortality exceeds baseline levels estimated using historical data. Rapid mortality surveillance, 'a system that generates daily or weekly counts of total mortality by age, sex, date of death, place of death' was proposed to inform decision-makers about the trajectory and magnitude of the pandemic, with excess mortality being the main focal point.[18]

The WHO defined death due to COVID-19 as 'A death resulting from a clinically compatible illness, in a probable or confirmed COVID-19 case, unless there is a clear alternative cause of death that cannot be related to COVID disease (e.g., trauma),' the death caused by COVID-19 may not be credited to another disease and should be counted freely of prior conditions that are associated with triggering a severe course of COVID-19.[19]

Saudi Arabia is divided into 13 regions covered by 20 health directorates. Each directorate has its autonomy in managing its health affairs according to policies and guidelines set by the Ministry of Health (MOH).[4] Epidemiological trends of any pandemic are essential to deciding whether to escalate or deescalate the precautionary measures and policies and allocate healthcare resources. Death per region has played a role in the lockdown for each area, and the identified risk factors are significant in prioritizing therapy and vaccination strategies. All COVID-19–related statistics including the number of SARS-CoV-2 polymerase chain reaction (PCR) tests, hospitalizations, intensive care unit (ICU) admissions, infected patients, and number of deaths are published and updated around the clock through a government website.[1] Our study aims to report the prevalence and epidemiological trends of mortality due to COVID-19 disease and report the most common comorbidities associated with mortality in KSA.

Methods

This study was performed following the Declaration of Helsinki and was approved by the central institutional review board of the Saudi MOH (log No: 20-168M). All methods were carried out under relevant guidelines and regulations of KSA. The institutional review board did not require informed consent, as this was not an interventional study, and the exemption was granted. Daily death data for the 20 health directorates were studied and obtained from the MOH and non-MOH public and private hospitals from March 23, 2020, through April 30, 2022. The analysis included all direct causes of death due to COVID-19. According to the WHO definition, the Saudi MOH registers COVID-19 as a possible cause of death under three categories. First direct death, in which death resulted from a clinical illness compatible with COVID-19 infection for confirmed or suspected cases. Second, indirect death resulted from another disease (e.g. diabetes, cardiovascular diseases, emphysema, etc.), and the patient's condition worsened due to a confirmed or suspected COVID-19. The last scenario is death unrelated to COVID-19, which is any death in a confirmed or suspected case with COVID-19, but COVID-19 was not a direct or indirect cause of death[1] (Fig. 1).

Since the beginning of the pandemic and to address any possible underestimation of reported COVID-19 mortality cases, the MOH has mandated that all healthcare facilities report all deaths to a mortality committee after filling a standardized form that contained all needed information to assess and evaluate the cause of death, including SARS-CoV-2 PCR test. The MOH established the COVID-19 mortality scientific committee to determine the causes of deaths during the COVID-19 pandemic, report COVID-19 mortalities, and evaluate the quality of care provided to patients with COVID-19. The committee met daily to assess and evaluate each case and decide whether mortality is direct, indirect, or unrelated to COVID-19. Any mortality case with confirmed or suspected COVID-19 was reported to the COVID-19 mortality scientific committee using the approved electronic form within 3 h from the time of death. These precautionary approaches are intended to enhance the accuracy of the number of reported cases and reduce the possibility of underestimation. The information in the form included demographics, symptoms, comorbidities, co-infections, vital signs, and laboratory values, all reported investigations, medications, ventilation setting and requirements, and complications during hospital stay (supplement 1). Analysis of all reports was sent to the statistics department, whereas only direct COVID-19–related mortalities were announced. By the end of 2020, the government of KSA approved the first COVID-19 vaccine Pfizer-BioNTech (Comirnaty), followed by the approval of the second vaccine in February 2021, Oxford-AstraZeneca (Vaxzevria), then Moderna (Spikevax). Vaccination was made available at many centers across KSA for Saudi and non-Saudi citizens at no cost.[20]

This report describes the prevalence of COVID-19 death in KSA, including mortality trends, COVID-19 vaccination, risk factors of

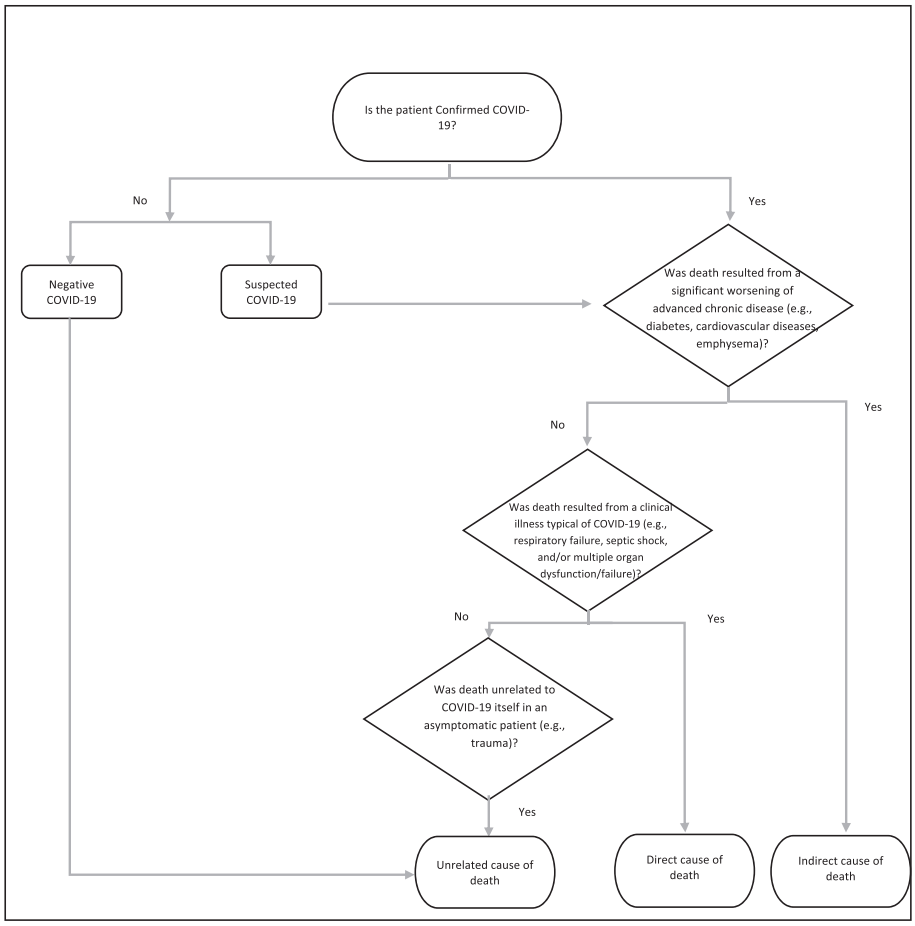


Fig. 1. Cause of death flowchart evaluation by Saudi MOH for COVID-19.

death, and population estimates for the prevalence of the various comorbidities. In addition, we report the association between comorbidities and death from COVID-19.

Statistical analysis

National and regional crude case fatality rate (cCFR) is defined as the number of mortality cases divided by the number of confirmed COVID-19 cases and the number of mortality cases per 100,000 population.[18] Descriptive statistics using numbers and percentages were used to describe the categorical variables such as nationality, age, gender, comorbidities, and geographical location of case fatalities. The median and interquartile (first quartile to third quartile; interquartile range [IQR]) for the time from the last COVID-19 vaccine till the time of death. Weekly mortality trends were reported as the number of deaths per period. The mortality trend was plotted per week and compared with international figures. Relative and absolute measures of association between comorbidities and death were calculated using odds ratios (ORs), attributable risk, and attributable risk percent. Population estimates for the prevalence of the various comorbidities were used to calculate population attributable risk estimates for comorbidities. Measures of association were derived from data for which each observation is classified along two dimensions, exposure and outcome, and each dimension is binary. The exposure is comorbidity (yes or no), and the outcome is death from COVID-19 (yes or no). Emphasis was placed on cause-specific death and death from COVID-19 vs death from another cause as the outcome of interest.

Comorbidities were analyzed individually and in combinations (e.g. diabetes and obesity together). Furthermore, the unexposed group was regarded in two different ways. First, we considered the unexposed to be those with no comorbidity and then investigated the extent to which the comorbidity (or combination of comorbidities) was associated with COVID-19 death vs no comorbidity. The second way to consider the unexposed was those without the comorbidity but possibly with various comorbidities. The former type of unexposed definition would reflect the isolated association of the comorbidity, whereas the latter would reflect the association of the comorbidity vs the set of other comorbidities seen in the population. To calculate the attributable risk and attributable risk percent associated with comorbidities, the risk of death from COVID-19 for the exposed and unexposed was calculated as follows[21]:

$$\text{Risk}_{\text{exposed}} = \frac{\text{Number of SARS} - \text{CoV2 with the comorbidity dying}}{\text{Total number of SARS} - \text{CoV2 with the comorbidity}}$$

and

$$\text{Risk}_{\text{unexposed}} = \frac{\text{Number of SARS} - \text{CoV2 without the comorbidity dying}}{\text{Number of SARS} - \text{CoV2 without the comorbidity}}.$$

Given these risk values, then the attributable risk and the attributable risk percent are as follows:

$$\text{Attributable risk} = \text{Risk}_{\text{exposed}} - \text{Risk}_{\text{unexposed}}$$

and

$$\text{Attributable risk percent} = \frac{\text{Risk}_{\text{exposed}} - \text{Risk}_{\text{unexposed}}}{\text{Risk}_{\text{exposed}}} \times 100.$$

Finally, the number needed to harm, defined as the number of subjects infected with SARS-CoV-2 who need to be exposed to comorbidity for one subject to die from COVID-19, was calculated as follows:

$$\text{NNH} = \frac{1}{\text{Attributable risk}}.$$

Measures of association were calculated together with 95% confidence intervals (CIs). Analyses were carried out using SAS/JMP, Version 15.0. SAS Institute Inc., Cary, NC, 1989–2021.

Results

Between March 23, 2020, and April 30, 2022, among the 10,110 subjects studied, 1025 had died from causes unrelated to COVID-19, and 9085 had died from causes directly related to COVID-19 from a total of 754,011 infected patients by COVID-19. The first reported death was on March 23 in Madinah province. Compared with cCFR in Brazil, China, Germany, Italy, South Korea, and the United States, KSA's cCFR was 1.21%. Among the 20 health directorates in KSA, Qurayyat Province accounted for the highest cCFR (4.70%), whereas Madinah reported a cCFR of 0.61% (Table 1). The national number of deaths per 100,000 population was 25.38, with Makkah province reporting the highest (48/100,000) and Tabuk reporting the lowest (12.18/100,000). The highest number of reported deaths was reported in the most populated provinces (Riyadh, Jeddah, Makkah, Jizan, Asir, and Hassa). The youngest reported death was a 42-day-old boy on August 18, 2020, who developed Multisystem Inflammatory Syndrome in Children, and the oldest was a male of 113 years. Males accounted for 71% of total deaths, and the highest percentage of deaths per age group was between the age of 60–69 years. The most common comorbidities associated with death were diabetes (60%), followed by hypertension (52%), obesity (12%), and heart failure (11%; Table 2). The highest number of reported mortalities was reached on week 15, then plateaued on week 16 and declined thereafter. Another peek was observed in week 63 with the second wave (Fig. 2).

Table 2
Demographics and baseline characteristics of deceased patients.

Demographics	Number of patients	Percentage
Male gender	6454	71%
Saudi nationality	4569	50%
Age ≥ 70 years	2487	27%
Medically free	1096	12%
Comorbidities		
Diabetic	5482	60%
Hypertension	4681	52%
Obesity	1058	12%
Heart failure	1002	11%
Renal impairment	776	9%
Respiratory disorders	849	9%
Neurological disorder	439	5%
Arrhythmias	324	4%
Thromboembolic disorders	286	3%
Thyroid dysfunction	305	3%
Liver impairment	82	1%

There was a total of 225 patients (2.8%) who received COVID-19 vaccine; of those, 213 (94.6%) got only one dose of the vaccine, 10 (4.4%) received two doses, and only two (0.8%) got the total three doses. This is in contrast to the proportion of the vaccinated individuals (25,082,132) in the community, which corresponds to 73.5% of the total population. The median time from the last vaccine until the day of death was 44.5 (IQR: 26.5–78) days. Those who got only one dose of the vaccine had a median time till death of 45 (IQR: 28–76) days, whereas patients who received two doses of vaccine had a median time of 113 (IQR: 20–177) days.

Across the whole cohort, there were 271 distinct comorbidity profiles. Diabetes (OR: 1.7, 95% CI 1.33–2.15) and hypertension (OR: 1.75, 95% CI 1.24–2.47) were the most common comorbidities associated with COVID-19 mortality. Diabetes, hypertension, arrhythmias (OR: 2.55, 95% CI 0.34–18.83), and obesity (OR: 3.79, 95% CI 2.01–7.13) were the top single comorbidity that was associated with COVID-19 mortality. Multiple coexisting comorbidities were compared with those without any comorbidity with regard to the risk of death, which showed that diabetes in combination with

Table 1
National crude case fatality rate and number of deaths per 100,000 population by April 30, 2022.

Health directorates	Population	Total cases	Total deceased	Number of deceased per 100,000 population	% Total deceased of cases
Riyadh	8,872,712	200,904	1590	17.92	0.79
Jeddah	5,031,820	86,766	1577	31.34	1.82
Makkah	2,512,462	44,006	1206	48.00	2.74
Jizan	1,670,569	30,465	691	41.36	2.27
Asir	1,940,123	44,094	681	35.10	1.54
Eastern	3,485,383	96,807	633	18.16	0.65
Al Hassa	1,305,172	45,631	613	46.97	1.34
Taif	1,387,686	25,872	429	30.91	1.66
Madinah	2,291,092	51,644	317	13.84	0.61
Qasim	1,520,434	28,074	304	19.99	1.08
Ha'il	746,046	15,611	230	30.83	1.47
Hafr Albatin	476,443	7587	141	29.59	1.86
Tabuk	968,414	11,949	118	12.18	0.99
Najran	621,040	13,135	116	18.68	0.88
Northern	390,656	6758	110	28.16	1.63
Al Baha	506,866	9152	102	20.12	1.11
Jouf	362,580	2906	81	22.34	2.79
Bisha	414,197	5745	59	14.24	1.03
Qunfotha	329,289	4306	45	13.67	1.05
Qurayat	180,430	894	42	23.28	4.70

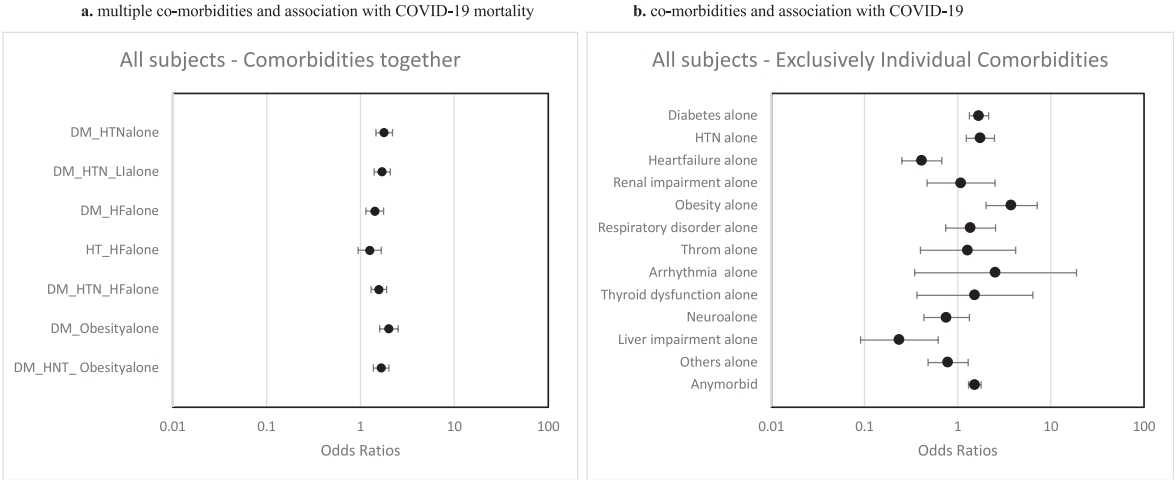
Fig. 2. Death per week trend by April 30, 2022.

hypertension and obesity (OR: 1.66, 95% CI 1.37–2.01) ranked as the highest multiple comorbidities associated with COVID-19 mortality, followed by the combination of diabetes with hypertension and heart failure (OR: 1.57, 95% CI 1.29–1.89), then diabetes and heart failure combined (OR: 1.42, 95% CI 1.14–1.76; Fig. 3). Table 3 shows the attributable risk of death and the number needed to harm as absolute measures of the risk of death for each comorbidity.

Discussion

This is the largest registry on the prevalence and epidemiological trends of COVID-19 mortality in KSA. Several countries have published the prevalence of mortalities related to COVID-19. The cCFR among European countries was the highest in Italy (2.8%).[22] Italy defines death related to COVID-19 in a broader term due to the lack of clear criteria for COVID-19 deaths. Death is recorded as related to COVID-19 in individuals who tested positive by real-time PCR for SARS-CoV-2 regardless of pre-existing conditions that might have contributed to their death.[13] In China, the case fatality rate reached 0.6% of confirmed cases, whereas in the United Kingdom, 0.8% of deaths were recorded as related to COVID-19.[22] In the United States, 81,863,725 COVID-19 cases were reported, with an estimated fatality rate of 1.2%.[22] The Centers for Disease Control and Prevention (CDC) reported a list of medical conditions that contributes to a severe COVID-19 outcome. Severity was defined as the need for hospitalization, ICU admission, intubation or mechanical ventilation, and death.[23]

The KSA has lower cCFR and deaths per 100,000 compared with many developed countries. This could be due to several factors, including multiple strategies of early preventive measures by the



diabetes (DM), hypertension (HT), heart failure (HF), renal impairment (RI), obesity (Ob), respiratory disorder (RD), thromboembolic disorder (TE), arrhythmia (Arr), thyroid dysfunction (TD), neurological disorder (ND), liver impairment (LI), some other comorbidity(ies) (Oth)

Fig. 3. Forest plots represented by odds ratios and confidence intervals.

Table 3
Comorbidities absolute measures of association with mortality.

Comorbidity	Exposed risk	Unexposed risk	Attributable risk	Attributable risk percent	NNH
Diabetes	0.96	0.84	0.11	11.94	8.76
Hypertension	0.96	0.86	0.10	10.34	10.09
Heart failure	0.97	0.90	0.06	6.69	15.49
Renal impairment	0.96	0.90	0.05	5.41	19.37
Obesity	0.98	0.90	0.09	8.68	11.71
Respiratory disorder	0.96	0.90	0.05	5.68	18.41
Thromboembolic disorders	0.96	0.91	0.05	5.26	19.90
Arrhythmias	0.96	0.91	0.06	5.94	17.48
Neurological disorder	0.93	0.91	0.02	2.21	48.72
Liver impairment	0.86	0.91	−0.05	−5.76	−20.24
Other	0.80	0.85	−0.05	−6.00	−21.30

NNH, number needed to harm.

government and early lockdown before the first COVID-19 mortality was reported. In addition, ICU bed expansion from 6360 beds to 10,401 beds (increased by 164%) during the pandemic period played a significant role in accommodating more critical patients. In addition, early outpatient intervention through the creation of 'fever clinics' in primary healthcare centers around the Kingdom. These clinics aimed to treat patients with early symptoms to prevent hospital admissions and decongest emergency departments in hospitals. Finally, MOH developed multiple protocols for COVID-19 that are continuously updated, which helped standardize the care across the Kingdom. More than 32 protocols covered different aspects of COVID-19-disease, including but not limited to COVID-19 treatment protocol, radiology preparedness, COVID-19 in pregnancy, ICU admission criteria, mechanical ventilation, and high flow nasal cannula and helmet use in COVID-19.

The number of cases in Italy had reached 16,798,998 in mid-2022.[22] In China, the prevalence of COVID-19 between men and women was also nearly the same (51.4% vs 48.6%). However, the fatality rate among men was significantly higher than women (relative risk (RR): 1.67, 95% CI 1.47–1.89; *P*-value: <0.001).[24,25] In the United Kingdom, mortality was also higher in men than women (OR: 1.47, 95% CI 1.26–1.73).[26] This is similar to our findings, as males accounted for more than two-thirds of mortalities. Different social and occupational factors could explain this among both genders.[27]

An increase in age was strongly associated with an increased risk of severe COVID-19 outcomes. One study reported case mortality and in-hospital mortality increased per age year by 3.4% and 7.4%, respectively (effect size case mortality: 1.074, 95% CI 1.061–1.087; effect size in-hospital mortality: 1.057, 95% CI 1.038–1.054).[28] Another study showed that age ≥50 years was an independent factor in ICU admission and in-hospital mortality.[29] According to the CDC, individuals aged ≥65 years accounted for 74.6% of the total COVID-19 deaths.[30] In the United Kingdom, age had the strongest impact on the COVID-19 associated mortality, people aged ≥50 years accounted for 98% of the total COVID-19 deaths, whereas 84% of the deaths were observed in people aged ≥70 years.[31] In Italy, the fatality rate was exceptionally high in the old population. A steep age-dependent fatality rate was seen in Bergamo province, 1.89% in the 70s, 4.84% in the 80s, and 11.06% in the age group above 90 years.[13] Our study showed that most mortalities occurred in patients aged 60–69 years.

Minorities ethnic groups have been especially affected by COVID-19, with an increase in fatality rate in the United States and the United Kingdom than in most populations.[32,33] Disparities in race, ethnicity, age, socio-economic status, and geographical location have been the key hallmarks of the previous US waves of COVID-19.

Individuals with pre-existing conditions are at an exceptionally high risk of death due to COVID-19.[14,34] In the United States, the most common comorbidities associated with COVID-19 mortality are influenza and pneumonia (48.9%), hypertension (18.6%), diabetes (15.4%), Alzheimer's disease, and other dementias (11.1%), and sepsis (9.8%).[30] After adjusting for sex, age, ethnicity, and race, it was found that having ≥3 underlying health conditions was associated with a greater risk of ICU admission (RR: 1.30, 95% CI 1.09–1.54) and death (RR: 1.81, 95% CI 1.44–2.28).[29] Diabetes has also been noted to increase the risk of death in both China and UK patients (RR: 4.43, 95% CI 3.49–5.61). $HbA_{1c} \geq 58$ mmol/mol had the greatest hazard ratio (HR): 1.95, 95% CI 1.83–2.07.[24,33] Our study suggests that diabetes is one of the important comorbidities associated with COVID-19 mortality. Knowing that diabetes prevalence in KSA is about 20%, this is particularly concerning. China and the United Kingdom have also established an association of increased risk of severe COVID-19 outcomes with respiratory disease (excluding asthma in both. RR: 3.43, 95% CI 2.42–4.87, HR 1.63, 1.55–1.71).[24,33] Severe asthma was associated with the risk of death in Chinese patients (asthma with oral corticosteroid use had the most significant HR: 1.13, 95% CI 1.01–1.26).[33] This is similar to our findings, as respiratory diseases were three times more likely to result in death due to COVID-19.

Vaccination was not available when our study began. Although this study was limited to reporting prevalence and did not explore vaccination in detail, another study conducted in KSA focused solely on vaccination launching and trends reported that one million doses were administered in about 3 months after the COVID-19 vaccine's introduction. The government made considerable efforts to make it widely available, and by May 2021, a total of 587 vaccine centers had been established, and more than 10 million doses had been administered.[20] Vaccines have been proven to be highly effective in reducing COVID-19 mortality ($P < 0.001$).[35] Our study showed that only 2.8% of patients who died due to COVID-19 received at least one dose of vaccine.

In our study, other pre-existing conditions that increased the risk of death in a patient with COVID-19 included cardiovascular diseases, renal impairment, obesity, thromboembolic disorders, thyroid dysfunction, and neurological disorders. A study from China found cardiovascular disease (RR: 6.75, 95% CI 5.40–8.43) and hypertension (RR: 4.48, 95% CI 3.69–5.45) among the most common comorbidities associated with COVID-19 mortality.[24] Another study in the United Kingdom showed that obesity was associated with an increased risk of death (BMI >40 kg/m^2; HR: 1.92, 95% CI 1.72–2.13). Chronic heart disease (HR: 1.17, 95% CI 1.12–1.22), liver disease (HR: 1.75, 95% CI 1.51–2.03), stroke/dementia (HR: 2.16, 95% CI 2.06–2.27), other neurological diseases (HR: 2.58, 95% CI 2.38–2.79), reduced kidney function (estimated glomerular filtration rate <30 had the greatest HR, HR: 2.52, 95% CI 2.33–2.72), autoimmune diseases (rheumatoid arthritis, lupus, or psoriasis, HR: 1.19, 95% CI 1.11–1.27), and other immunosuppressive

conditions (HR: 1.70, 95% CI 1.34–2.16), were also associated with increased risk of death.[33]

It is important to note that we reported the association of different comorbidities, which does not necessarily mean correlation or causality. Furthermore, two critical assumptions are necessary for the validity of the measures of association in this study. First, it must be assumed that the presence or absence of comorbidity is independent of the likelihood of getting infected by the COVID-19 virus. Second, the distribution of the comorbidities among the study subjects who died from unrelated causes matches the distribution of the comorbidities in the population of interest.

Although the study illustrated the impact of different comorbidities on the rate of COVID-19–related mortality, some limitations must be acknowledged. First, confounding factors that may contribute to the disease severity, such as air pollution, humidity, and temperature were not investigated. However, each mortality case in this study was subjected to a scientific committee evaluation that determined the probability of a COVID-19–related event. In addition, because vaccinations were introduced 9 months after the study began, the vaccination effect on COVID-19 mortality may not have been accurate. It is important to know that we did not perform a quantitative comparison between KSA and G20 countries based on well-recognized epidemiological models, as this was not the aim of the study. Direct comparison based on semiquantitative comparison should be interpreted with caution.

Conclusion

CFR and death per 100,000 population in KSA are among the lowest worldwide due to multiple factors. Several comorbidities have been identified, namely, diabetes, hypertension, obesity, and cardiac arrhythmias. Therefore, understanding COVID-19 mortality epidemiological trends and associated factors will help prepare healthcare systems for inevitable pandemics, hence reducing the healthcare burden and expediting recovery. Data from this study could assist in public health decision-making and preparedness efforts.

Author statements

Acknowledgments

The authors would like to thank Tahani Altamimi, Lama Alrabah, Fouziah Alharbi, Mohammed Masaad Al-otaibi, Razan Zahran, and Khaulah Alokili for their assistance with collecting data.

Ethical approval

This study was performed in accordance with the Declaration of Helsinki and was approved by the central institutional review board of the Saudi Ministry of Health (log No: 20-168M). The institutional review board did not require informed consent as this was not an interventional study, and the exemption was granted.

Funding

This research project was not funded by any commercial or non-commercial organization.

Competing interests

The authors of this article have no conflicts of interest or any direct or indirect financial interests.

Consent for publication

Approval for publication was granted by the central institutional review board of the Saudi Ministry of Health.

Availability of data and materials

All data and materials are available on request.

Author contributions

D.A.A. was the primary investigator and contributed to writing the proposal, figures, study design, data collection, data analysis, data interpretation, and writing the article. W.A. and R.A. contributed to literature search and writing the article. H.Y.A., Z.A., F.A., H.E., A.A.M., M.M.A., E.A., T.A., M.S.A., M.S.B., A.A.A., M.H.A., A.A., and R.N. contributed to reviewing article, reviewing cases, and methodology. E.B.D. and G.M. contributed to data analysis. A.H.A. contributed to study design, data interpretation, and writing the article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2022.07.014.

References

1. COVID 19 Dashboard: Saudi Arabia. Retrieved 30 April 2022, from, https://covid19.moh.gov.sa/; 2021.
2. *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*. 2021. Retrieved 31 December 2021, from, https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020.
3. WHO Coronavirus (COVID-19) Dashboard. Retrieved 30 April 2022, from, https://covid19.who.int/; 2021.
4. Saudi MOH. *Statistical Yearbook*. 2021. Retrieved 30 April 2022, from, https://www.moh.gov.sa/en/Ministry/Statistics/book/Documents/book-Statistics.pdf.
5. Influenza (Seasonal). Retrieved 30 April 2022, from, https://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal); 2021.
6. McMullen P, Pytel P, Snyder A, Smith H, Vickery J, Brainer J, et al. A series of COVID-19 autopsies with clinical and pathologic comparisons to both seasonal and pandemic influenza. *J Pathol Clin Res* 2021;**7**(5):459–70. https://doi.org/10.1002/cjp2.220.
7. *Leading causes of death US 2020-2021 | Statista*. 2021. Retrieved 31 December 2021, from, https://www.statista.com/statistics/1254560/leading-causes-of-death-in-the-us-average-number-daily/.
8. Lau H, Khosrawipour T, Kocbach P, Ichii H, Bania J, Khosrawipour V. Evaluating the massive underreporting and undertesting of COVID-19 cases in multiple global epicenters. *Pulmonology* 2021;**27**(2):110–5.
9. Baggiani A, Briani S, Luchini G, Giraldi M, Martino MC, Porretta A, et al. Management of healthcare areas for the prevention of COVID-19 emergency in an Italian teaching hospital in Pisa, Tuscany: a hospital renovation plan. *Infect Control Hosp Epidemiol* 2020;**41**(11):1368–9.
10. Maniscalco P, Poggiali E, Quattrini F, Ciatti C, Magnacavallo A, Caprioli S, et al. The deep impact of novel CoVID-19 infection in an Orthopedics and Traumatology Department: the experience of the Piacenza Hospital. *Acta Biomed* 2020;**91**(2):97.
11. Torres I, Sippy R, Sacoto F. Assessing critical gaps in COVID-19 testing capacity: the case of delayed results in Ecuador. *BMC Public Health* 2021;**21**(1):1–8.
12. Itamura K, Wu A, Illing E, Ting J, Higgins T. Youtube videos demonstrating the nasopharyngeal swab technique for SARS-CoV-2 specimen collection: content analysis. *JMIR Public Health Surveill* 2021;**7**(1):e24220.
13. Modi C, Böhm V, Ferraro S, Stein G, Seljak U. Estimating COVID-19 mortality in Italy early in the COVID-19 pandemic. *Nat Commun* 2021;**12**(1):1–9.
14. Soto A, Quiñones-Laveriano DM, Azañero J, Chumpitaz R, Claros J, Salazar L, et al. Mortality and associated risk factors in patients hospitalized due to COVID-19 in a Peruvian reference hospital. *PLoS One* 2022;**17**(3):e0264789.
15. Wu X, Nethery RC, Sabath MB, Braun D, Dominici F. Air pollution and COVID-19 mortality in the United States: strengths and limitations of an ecological regression analysis. *Sci Adv* 2020;**6**(45):eabd4049.
16. Chen S, Prettner K, Kuhn M, Geldsetzer P, Wang C, Bärnighausen T, et al. Climate and the spread of COVID-19. *Sci Rep* 2021;**11**(1):1–6.
17. Ganegoda NC, Wijaya KP, Amadi M, Erandi KKW, Aldila D. Interrelationship between daily COVID-19 cases and average temperature as well as relative humidity in Germany. *Sci Rep* 2021;**11**(1):1–16.
18. *Revealing the toll of COVID-19*. 2021. Retrieved 30 April 2022, from, https://www.who.int/publications/i/item/revealing-the-toll-of-covid-19.

19. *Estimating mortality from COVID-19*. 2021. Retrieved 30 April 2022, from, https://www.who.int/news-room/commentaries/detail/estimating-mortality-from-covid-19.
20. Assiri A, Al-Tawfiq JA, Alkhalifa M, Al Duhailan H, Al Qahtani S, Dawas RA, et al. Launching COVID-19 vaccination in Saudi Arabia: lessons learned, and the way forward. *Travel Med Infect Dis* 2021:102119.
21. Population Attributable Risk (PAR). Retrieved 30 April 2022, from, https://link.springer.com/referenceworkentry/10.1007%2F978-1-4020-5614-7_2685; 2021.
22. Mortality Analyses - Johns Hopkins Coronavirus Resource Center. Retrieved 10 May 2022, from https://coronavirus.jhu.edu/data/mortality, ; 2021.
23. *Underlying medical conditions associated with higher risk for severe COVID-19: information for healthcare providers*. 2021. Retrieved 30 April 2022, from, https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-care/underlyingconditions.html.
24. Deng G, Yin M, Chen X, Zeng F. Clinical determinants for fatality of 44,672 patients with COVID-19. *Crit Care* 2020;**24**.
25. Li G, Liu Y, Jing X, Wang Y, Miao M, Tao L, et al. Mortality risk of COVID-19 in elderly males with comorbidities: a multi-country study. *Aging (Albany NY)* 2021;**13**(1):27.
26. Elliott J, Bodinier B, Whitaker M, Delpierre C, Vermeulen R, Tzoulaki I, et al. COVID-19 mortality in the UK Biobank cohort: revisiting and evaluating risk factors. *Eur J Epidemiol* 2021;**36**(3):299–309.
27. Griffith DM, Sharma G, Holliday CS, Enyia OK, Valliere M, Semlow AR, et al. Men and COVID-19: a biopsychosocial approach to understanding sex differences in mortality and recommendations for practice and policy interventions. *Prev Chronic Dis* 2020;**17**:200247. https://doi.org/10.5888/pcd17.200247external icon.
28. Starke KR, Reissig D, Petereit-Haack G, Schmauder S, Nienhaus A, Seidler A. The isolated effect of age on the risk of COVID-19 severe outcomes: a systematic review with meta-analysis. *medRxiv* 2021.
29. Kim L, Garg S, O'Halloran A, Whitaker M, Pham H, Anderson EJ, et al. Risk factors for intensive care unit admission and in-hospital mortality among hospitalized adults identified through the US coronavirus disease 2019 (COVID-19)-associated hospitalization surveillance network (COVID-NET). *Clin Infect Dis* 2021;**72**(9):e206–14.
30. NVSS - Provisional Death Counts for COVID-19. Retrieved 2 January 2022, from, https://www.cdc.gov/nchs/covid19/mortality-overview.htm; 2022.
31. Merchant HA, Kow CS, Hasan SS. COVID-19 first anniversary review of cases, hospitalization, and mortality in the UK. *Expet Rev Respir Med* 2021:1–6.
32. Zelner J, Trangucci R, Naraharisetti R, Cao A, Malosh R, Broen K, et al. Racial disparities in coronavirus disease 2019 (COVID-19) mortality are driven by unequal infection risks. *Clin Infect Dis* 2021;**72**(5):e88–95.
33. Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al. OpenSAFELY: factors associated with COVID-19 death in 17 million patients. *Nature* 2020;**584**(7821):430–6. https://doi.org/10.1038/s41586-020-2521-4.
34. Bergman J, Ballin M, Nordström A, Nordström P. Risk factors for COVID-19 diagnosis, hospitalization, and subsequent all-cause mortality in Sweden: a nationwide study. *Eur J Epidemiol* 2021;**36**(3):287–98.
35. Jabłońska K, Aballéa S, Toumi M. The real-life impact of vaccination on COVID-19 mortality in Europe and Israel. *Public Health* 2021;**198**:230–7.

Public Health 215 (2023) 12–16

Contents lists available at ScienceDirect

Public Health

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

Themed Paper – Original Research

Strengthening the equity focus of applied public health research: introducing the FOR EQUITY platform

J. Popay [a, *], C.K. Chekar [a], A. Griffiths [a], E. Halliday [a], H. Kaloudis [a], R. Leiper [a], K. Panagaki [a], A. Porroche-Escudero [b]

[a] Division of Health Research, Faculty of Health & Medicine, Lancaster University, Bailrigg, Lancaster, LA1 4YG, UK
[b] Lancaster Environment Centre, Lancaster University, Bailrigg, Lancaster LA1 4YE, UK

ARTICLE INFO

Article history:
Received 1 June 2022
Received in revised form
21 November 2022
Accepted 25 November 2022
Available online 4 January 2023

Keywords:
Health equity
Health inequalities
Intersectionality
Toolkits

ABSTRACT

Objectives: Much applied health research pays insufficient attention to potential unequal impacts across social groups or is typically focused on a single dimension (e.g. socio-economic status), rarely considering the intersecting social processes driving inequalities (e.g. racism, sexism, classism). All health research needs a strong intersectional equity focus in order to inform action to reduce health inequalities as well as improve population health.

Study design: Focus On Research and Equity (*FOR EQUITY*) is a new Web-based platform aiming to strengthen the intersectional equity focus of applied health research.

Methods: The platform was developed in collaboration with members of the public, practitioners and researchers working internationally. The development involved a systematic review of academic and grey literature, a series of workshops and user testing.

Results: FOR EQUITY encompasses (1) a Health Inequalities Assessment Tool, with an intersectional perspective on inequalities; (2) a FOR EQUITY Guidance Inventory providing access to a range of international research toolkits and guidance; and (3) a FOR EQUITY Library including case studies illustrating how researchers have attempted to integrate an equity lens into the research process and more general resources on health inequalities.

Conclusion: FOR EQUITY can support researchers to strengthen the equity lens in their studies to make research evidence more relevant for action to reduce social and health inequalities. However, a single focus on toolkits is unlikely to sufficiently address the barriers to embedding equity in research. A mainstreaming strategy to transform the very roots of the 'institution of research' is required.

Introduction: the problem

Persistent inequalities in a wide range of health outcomes across social groups and places are "a stark reminder of the unequal nature of our societies" (p. 1).[1] These inequalities are systematic, avoidable through policy action, and therefore unfair.[2,3] Globally, the health research community has produced a very significant body of high-quality research on the *nature and causes* of these inequalities, with a smaller body of work evaluating the impact of policies and other initiatives on reducing inequalities. In a recent paper, Kelly–Irving et al. argue that the ongoing debate as to whether this body of research has demonstrated a causal relationship between health inequalities and inequalities in the resources and power people have available to them poses major challenges for policy-makers. But other equally important challenges for action to reduce health inequalities emerge from the current body of health research.

First, research focused explicitly on understanding health inequalities or informing action to reduce them constitutes only a small part of the health research endeavour in the United Kingdom and globally. The bulk of this research is focused on understanding drivers of population health outcomes and/or evaluating policies, interventions and other initiatives aimed at improving population health. Much of this research fails either in part or whole to consider the differential impact of the factors and/or processes driving population health or the distributional impacts across social groups of policies and other actions.[4–6] As a result, much health research is unable to inform the design of policies and other actions

* Corresponding author.
E-mail address: j.popay@lancaster.ac.uk (J. Popay).

https://doi.org/10.1016/j.puhe.2022.11.018

that have potential to reduce health inequalities at the same time as improving population health. For example, an evaluation of the progress made in mainstreaming an equity focus within a regional research collaboration found that the lack of an explicit strategy, compounded by different understandings of health inequalities, constrained the ability of partner organisations and their staff to engage in and/or design and deliver health equity sensitive research.[5]

Second, research that is *directly* focused on understanding the causes of health inequalities and/or evaluating the effectiveness of action to reduce inequalities has tended to focus on a single dimension of these inequalities. In the United Kingdom and Europe, for example, much of the research has focused on health inequalities across socio-economic groups; in the USA, much of the research has focused on health inequalities between ethnic/racial groups; and in New Zealand, much of the research has focused on inequalities between indigenous people and other New Zealanders. Studies that consider the intersection between a broader range of social identities/categories (e.g. class, race and gender) and social processes driving inequalities (racism and sexism) are less common. As Holman and Walker argue, by focusing only on specific characteristics in isolation, research can 'not match the reality that people simultaneously embody multiple characteristics and are therefore potentially subject to multiple forms of discrimination' (p. 239).[7]

Finally, applied non-clinical health research continues to be weighted to the design and evaluations of lifestyle and behavioural interventions that typically fail to take into account the influence of wider social determinants (e.g. poverty, insecure housing and employment). As a result, although they can improve population health, these interventions have limited ability to reduce health inequalities, and they may even increase inequalities because they are more effective with more 'advantaged' groups.[8] For example, in their review of studies of behavioural interventions, Alcantara et al. concluded that "45% of studies tested social determinant of health moderators, and yet, most studies were atheoretical, lacked sample and lifecourse diversity, were of poor methodological quality, and were focused at the individual-level" (p. 141).[9]

Both research and policy interest in health inequalities have grown in recent years, with the publication of major national and international reports, notably the 2008 report of the WHO sponsored Commission on the Social Determinants of Health. In this context, a plethora of resources, such as guidance and toolkits, have been produced to support the design of policies and practices that simultaneously improve population health and reduce health inequalities. International repositories have also been established to help people navigate the materials available.[10] There are some similar resources aimed at strengthening the equity lens in research. For example, in response to the findings of their review, Alcantara et al. developed a framework to support researchers to integrate a focus on social determinants of health into the design and evaluation of behavioural interventions to "centre health equity, social justice, and community health" (p. 141).[9] There was a similar initiative in relation to dental health promotion interventions in 2012.[11] Generally, however, although increasing in number, resources targeting support for the development of more equity sensitive applied health research are much less common than those designed for policy-makers and practitioners.[12] In addition, no single repository exists for the resources that do exist.

The purpose and development of FOR EQUITY

Focus On Research and Equity (*FOR EQUITY:* https://forequity.uk) is a new Web-based platform that aims to support the integration of an 'intersectional equity lens' in all types of applied health research. Given the widespread inequalities in health-related outcomes to be found in all countries, strengthening the equity focus of any research, whatever its focus and design, will add valuable dimensions to the evidence base to inform more effective action to reduce health inequalities. For example, hypertension is a major public health issue. The condition significantly increases the risks of heart, brain, kidney and other diseases and is a major cause of premature death worldwide. Yet, globally an estimated 46% of adults with hypertension are unaware that they have the condition; only 42% of cases are diagnosed and treated; and only one in five adults (21%) with hypertension have it under control. Crucially, the incidence of hypertension amongst people in the most disadvantaged areas in the United Kingdom is double that of those in the most advantaged areas.[13] Yet, despite considerable investment in research questions of equity around unmanaged hypertension across social groups and the impact of social determinants on help seeking behaviour and adherence to treatment are not systematically considered in studies. As a result, much research evidence cannot inform the development of more effective approaches to reduce *inequalities* in the diagnosis and treatment of hypertension and could lead to interventions that actually widen these inequalities.

FOR EQUITY is a new online tool designed to enable its users to navigate the complex and expanding landscape of resources, guidance and toolkits that are being produced to strengthen an intersectional equity focus in research. Over time, FOR EQUITY could contribute to an increase in the volume of applied research directly focused on the social determinants of health inequalities, but this is not its primary purpose. Rather it aims to help to increase the relevance of all health research to action that can reduce health inequalities, regardless of the focus and study design.

The platform was developed in collaboration with members of the public and an expert advisory group of practitioners and academics from the public and third sector working in the United Kingdom and internationally. It was jointly funded by two national research collaborations in England: the National Institute for Health and Care Research School for Public Health Research (NIHR SPHR) and NIHR Applied Research Collaboration North West Coast (NIHR ARC NWC).

The resource was tested with potential users in a number of ways at different time points in the development. Early in the development phase in 2020, two workshops were held to identify options for introducing the concept of intersectionality into the resource. A range of academic and practitioner expertise on intersectionality was represented. Later in the development phase in 2021, a prototype resource was then tested with three teams at different stages of the research process: (1) a pilot inquiry into voluntary service delivery during the COVID-19 crisis, (2) a large UKRI transdisciplinary multiuniversity project on green and blue spaces, and (3) a large NERC funding bid on air inequality. The utility and accessibility of the material were discussed with these teams. Following this initial stage, the resource was revised, and its utility was tested in two further workshops between April and June 2021. These aimed to test the revised resource with frequently used research designs: evidence reviews, mixed methods study designs; and use of secondary/routine data. A total of 21 people participated in the workshops, which were held online and lasted up to 3 h. Workshop discussions were audio recorded, and participants were asked to complete a written consent form.

Finally, between November and December 2021 (3 months before FOR EQUITY was launched), 21 researchers with diverse disciplinary and methodological expertise working in the NIHR ARC NWC and SPHR (the funders of FOR EQUITY) and two senior staff of third sector organisations involved in research were asked via email to explore the resource and provide feedback. Some of this

group had attended the earlier workshops, but none had been involved directly in developing the resource. They received a one-page introductory leaflet to help them navigate the site, and they were told that the overall architecture could not be changed. In this context, they were asked to provide general feedback and to comment on four issues in particular: the ease of navigating through the site; clarity of the purpose of the resources and content of its parts; potential problems with the content (e.g. gaps in coverage); and format and content of the case studies in the FOR EQUITY library (see below).

We received detailed feedback via these routes, much of it very positive. Improvements resulting from these activities included making navigation of the site easier by rearranging content on the landing page, redrafting some of the text to simplify it, correcting grammatical errors and inconsistent use of language and repairing broken links.

The FOR EQUITY platform

FOR EQUITY has three components (see Fig. 1).

- *The Health Inequalities Assessment Tool* (*HIAT*) supports people to reflect on how they can better integrate an equity lens throughout the research process from identifying issues to study and formulating the questions to be addressed through study design and conduct to dissemination and knowledge exchange. It is not a 'how to' manual but rather a space in which reflexive questions are provided to encourage people to think about the most appropriate way for intersectional equity issues to be integrated into their research. There is a strong focus on how involving people with lived experience and policy or practice expertise can help strengthen the equity lens in research. HIAT was originally developed by the NIHR Collaboration for Applied Research and Care in the North West Coast (now ARC NWC).[12] It has now been revised to include a broader intersectional perspective on inequalities. Users are supported through five linked sections (see Fig. 2). Four relate to the research process: mapping health inequalities relevant to your research; integrating equity into your research questions; designing and conducting equity sensitive research; and prioritising equity relevant findings in dissemination. The fifth considers principles of equity sensitive research practice highlighting researchers' responsibilities to deliver evidence that can contribute to reduce health inequalities and the social inequalities that drive these.
- The FOR EQUITY *Guidance Inventory* is based on a scoping review of published tools/guides that provide practical advice on how to embed different dimensions of equity into research. The review objectives were to (1) scope the resources' intended use in the research process; (2) compare the theoretical orientation of resources; (3) consider the coverage of public and practitioner involvement in the resources; and (4) identify the extent of the resources' application and evaluation in practice. Searches were conducted of key academic databases and grey literature published in English, and topic experts were also consulted. A total

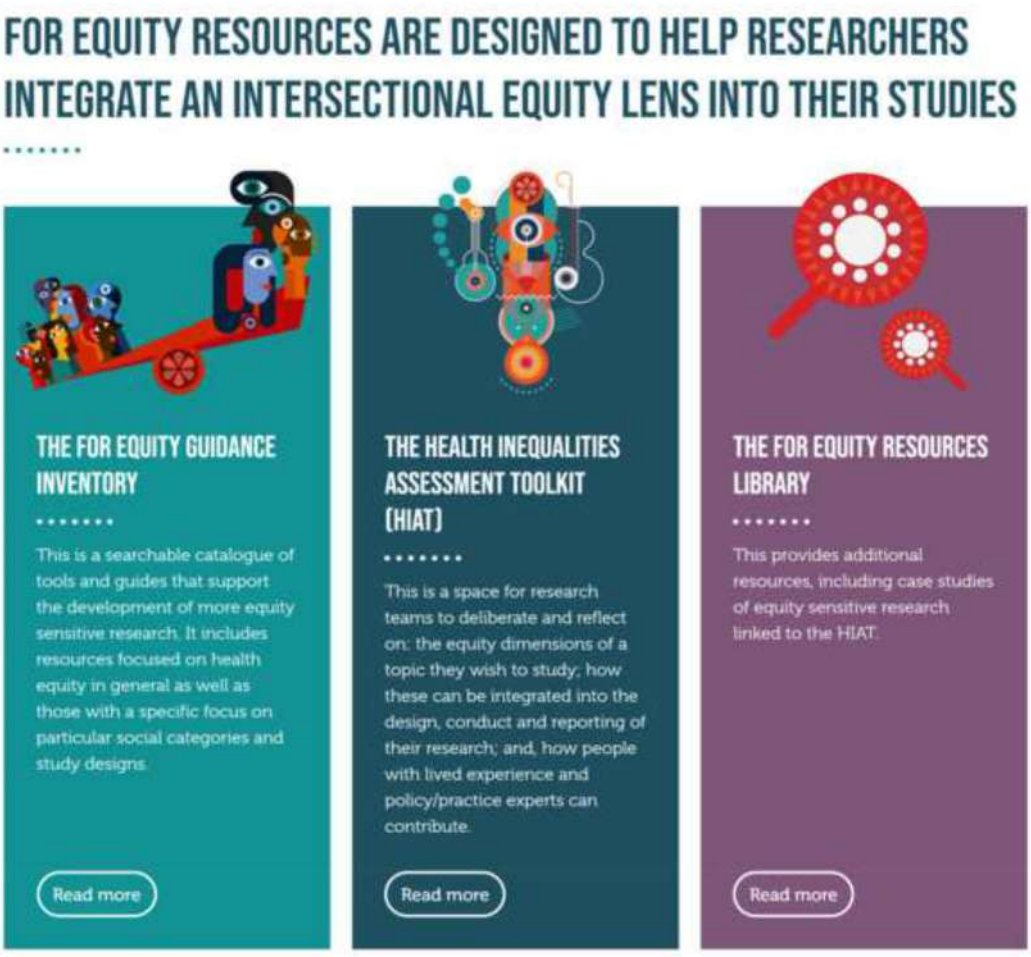


Fig. 1. FOR EQUITY about here.

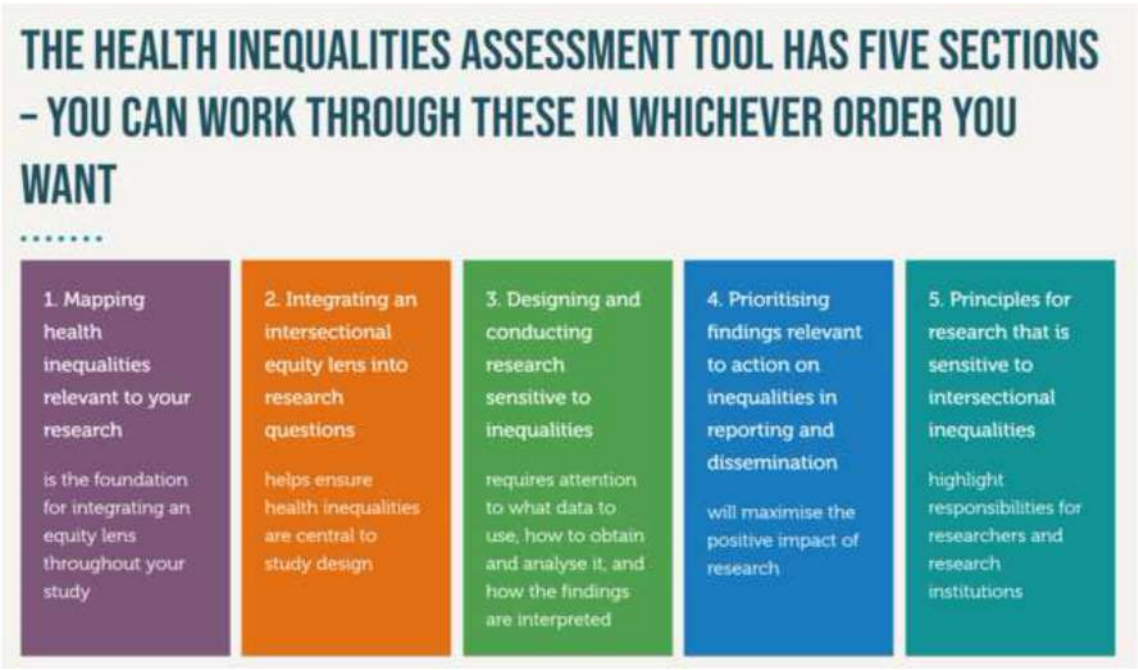


Fig. 2. HIAT about here.

of 2622 records were identified and screened, leading to the identification of the 22 resources included in the FOR EQUITY Inventory. Each entry in the Inventory is summarised outlining the resource's intended use in the research process and its theoretical orientation plus evidence of application in other research. The Inventory also includes resources aiding understanding about how discrimination and inequalities intersect across specific social categories, including social class/socioeconomic status, ethnicity,[14] gender or sex[15] and disability.[16] The inventory can be searched by topic or using free text searches. Half the resources in the Inventory were developed in Canada and USA, with some from the United Kingdom and New Zealand and only one from a low-income country: South Africa. Resources comprised a range of formats, including frameworks, reporting guidelines or checklists, principles, evaluation tools and a decision-making aid. Some focus on a particular methodology (e.g. randomised controlled trials or systematic reviews). Resources also varied in the extent that these were process orientated (e.g. checklists to aid reporting) or served as a more reflective tool, for example, to encourage researchers to explore particular equity dimensions.

- The *FOR EQUITY Library* provides a series of case studies linked to the HIAT that illustrate how researchers have attempted to integrate an equity lens into the research process. There are also general resources, including Web sites, papers and reports covering a range of issues involved in developing, delivering and disseminating research that is equity sensitive.

Adopting an intersectional lens

Critically, FOR EQUITY resources are designed to help researchers integrate an intersectional equity lens into their studies. Most resources identified in our review acknowledged multidimensional and structural causes of health inequities. However, they differed in the extent that social categories and processes were conceptualised as singular or interacting factors shaping unequal outcomes and experiences. Some resources signposted users to PROGRESS-Plus as a tool to encourage consideration of the range of characteristics shaping health inequity.[17] However, most resources referring to intersectionality considered this in the context of a particular social category (ethnicity or race, gender/sex, disability). These resources were typically underpinned by a concern about structural drivers experienced by groups sharing particular characteristics and encouraged greater consideration of these.

In the context of these findings, FOR EQUITY includes reflective questions encouraging research teams to consider the relevance of an intersectional perspective to their studies in terms of the social groups studies focus on, the nature of inequalities in the experience of these groups and the factors that may individually or in interaction, have caused these unequal experiences or impacts. The FOR EQUITY library also provides links to a range of resources for users wanting to explore intersectionality in more depth (including videos, journal articles and Web sites).

A final word of caution

FOR EQUITY aims to strengthen the equity focus in applied health research. We anticipate that this new platform will contribute to improving researchers' awareness and understanding of the intersectional dimensions of health inequalities and the salience of the wider social determinants of these inequalities, whatever the focus and design of their study. Early evaluation of the HIAT, for example, found that the tool increased understanding among researchers about the drivers of inequalities and a greater appreciation that addressing health inequalities is 'everyone's business' (p. 569).[12] Over time, such changes in awareness and behaviours could, in turn, contribute to a greater body of research on the distributional impacts of public health policies and interventions to inform policy and practice decision-making.

Nevertheless, a single focus on toolkits and guidance is likely to be insufficient to address the barriers to routinely embedding equity in research. One risk is that their use could serve as a box-ticking exercise without genuine consideration about the

implications of health and social inequalities in the research process, including the role of wider social determinants of health in (re)producing these inequalities. In fact, existing efforts to promote health equity within research organisations have suggested the need to move beyond limited technical approaches: primarily capacity building initiatives and tool usage. This will require a mainstreaming agenda aimed at transforming the very roots of the 'institution of research', including researchers, research organisations and research funding bodies.[10]

Author statements

Acknowledgements

The authors would like to acknowledge the contribution of other SPHR member collaborators based at FUSE (Clare Bambra and Vicki McGowan), SCHARR (Dan Holman and Sarah Salway) and LiLaC (Rebecca Mead and Margaret Whitehead) as well as our practice collaborator (John Hume, People's Health Trust) involved in the development of FOR EQUITY. The authors would also like to thank all participants attending project workshops during 2021 who gave their valuable time and insights to support revisions to the HIAT tool.

Ethical approval

None declared.

Funding

This project was funded by the National Institute for Health and Care Research (NIHR) School for Public Health Research (SPHR), Grant Reference Number PD-SPH-2015 and the NIHR ARC Applied Research Collaboration North West Coast, Grant Reference Number NIHR200182. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Competing interests

None declared.

References

1. Kelly-Irving M, Ball WP, Bambra C, Delpierre C, Dundas R, Lynch J, et al. Falling down the rabbit hole? Methodological, conceptual and policy issues in current health inequalities research. *Crit Publ Health* 2022:1–11.
2. McCartney G, Popham F, McMaster R, Cumbers A. Defining health and health inequalities. *Publ Health* 2019;**172**:22–30.
3. Whitehead M. A typology of actions to tackle social inequalities in health. *J Epidemiol Community Health* 2007;**61**:473–8.
4. Porroche-Escudero A, Popay J. If I can't dance it is not my revolution: impactful health promotion research needs an equity lens. In: Potvin L, Jourdan D, editors. *Handbook of health promotion research*. Canada: Springer; 2022. p. 693–704.
5. Porroche-Escudero A, Popay J, Ward F, Ahmed S, Akeju D, Cloke J, et al. From fringe to centre-stage: experiences of mainstreaming health equity in a health research collaboration. *Health Res Pol Syst* 2021;**19**:1–14.
6. Lorenc T, Petticrew M, Welch V, Tugwell P. What types of interventions generate inequalities? Evidence from systematic reviews. *J Epidemiol Community Health* 2013;**67**:190–3.
7. Holman D, Walker A. Understanding unequal ageing: towards a synthesis of intersectionality and life course analyses. *Eur J Ageing* 2021;**18**:239–55.
8. McGill R, Anwar E, Orton L, Bromley H, Lloyd-Williams F, Oflaherty M, et al. Are interventions to promote healthy eating equally effective for all? Systematic review of socioeconomic inequalities in impact. *BMC Publ Health* 2015;**15**:457.
9. Alcántara C, Diaz SV, Cosenzo LG, Loucks EB, Penedo FJ, Williams NJ. Social determinants as moderators of the effectiveness of health behavior change interventions: scientific gaps and opportunities. *Health Psychol Rev* 2020;**14**:132–44.
10. Pauly B, Martin W, Perkin K, van Roode T, Kwan A, Patterson T, et al. Critical considerations for the practical utility of health equity tools: a concept mapping study. *Int J Equity Health* 2018;**17**:48.
11. Watt RG, Sheiham A. Integrating the common risk factor approach into a social determinants framework. *Community Dent Oral Epidemiol* 2012;**40**:289–96.
12. Porroche-Escudero A, Popay J. The Health Inequalities Assessment Toolkit: supporting integration of equity into applied health research. *J Publ Health* 2020;**43**:567–72.
13. NHS England. Evidence for the 5 clinical priorities. Available from:: https://www.england.nhs.uk/about/equality/equality-hub/national-healthcare-inequalities-improvement-programme/core20plus5/evidence-for-the-five-clinical-priorities/ [accessed 20 November 2022]. [date?].
14. Salway Sarah, Barley R, Allmark P, Gerrish K, Higginbottom G, Johnson M, et al. Enhancing the quality of published research on ethnicity and health: is journal guidance feasible and useful? *Divers Health Care* 2011;**8**:155–65.
15. Schiebinger L, Klinge I, Sánchez de Madariaga I, Paik H, Schraudner M, Stefanick M. *Gendered innovations in science, health & medicine, engineering and environment*. 2011-2021 [26 April 2021]; Available from: http://genderedinnovations.stanford.edu/index.html.
16. Berghs M, Atkin K, Graham H, Hatton C, Thomas C. *Implications for public health research of models and theories of disability: a scoping study and evidence synthesis*. 2016. Southampton (UK).
17. O'Neill J, Tabish H, Welch V, Petticrew M, Pottie K, Clarke M, et al. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *J Clin Epidemiol* 2014;**67**:56–64.

Public Health 215 (2023) 27–30

Contents lists available at ScienceDirect

Public Health

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

Themed Paper — Original Research

The association of body mass index and odds of type 2 diabetes mellitus varies by race/ethnicity☆

S. Strings [a,*], C. Wells [b], C. Bell [c], A.J. Tomiyama [d]

[a] Department of Sociology, 3151 Social Sciences Plaza A, University of California, Irvine, Irvine CA 92697, USA
[b] Advanced Research Computing, University of California, Los Angeles, Los Angeles CA, USA
[c] Social, Behavioral, and Population Sciences, School of Public Health & Tropical Medicine, Tulane University, New Orleans LA, USA
[d] Department of Psychology, University of California, Los Angeles, Los Angeles CA, USA

ARTICLE INFO

Article history:
Received 31 May 2022
Received in revised form
14 November 2022
Accepted 25 November 2022
Available online 10 January 2023

Keywords:
Diabetes
Obesity
Race/ethnicity

ABSTRACT

Objective: This study aimed to examine the association between body mass index (BMI; weight [kilogram]/height^2 [meter]) and type 2 diabetes mellitus (T2DM) among the largest three largest racial/ethnic groups in the United States.

Methods: We compiled 10 waves of the continuous National Health and Nutrition Examination Survey from 1999–2000 through 2017–2018. Participants ($N = 45{,}514$) were those who had data on BMI, HbA_{1c}, and demographics. We estimated associations between BMI and prediabetes/T2DM odds for Black, Latine, and White participants.

Results: BMI was associated with 10% higher odds of prediabetes/T2DM vs. having normal HbA_{1c} levels (odds ratio = 1.10, 95% confidence interval = 1.10–1.11) for Latine and White individuals. However, the association between BMI and prediabetes/T2DM was significantly weaker among Black individuals. When focusing on T2DM prevalence, the association with BMI for Black participants was even weaker (odds ratio = 0.97, 95% confidence interval = 0.95–0.98).

Conclusions: The unstable associations between BMI and T2DM across race indicate that BMI has received unwarranted focus as a prime predictor of T2DM. Relying on BMI introduces bias in T2DM risk estimations especially in Black individuals. Focusing on BMI places the onus on individuals to change and increases weight stigma, which can worsen health outcomes. Instead, policymakers should focus on social determinants of T2DM and its concomitant racial/ethnic disparities.

Introduction

Rising rates of what is defined as "obesity" (body mass index [BMI] ≥30 kg/m^2) have been identified as a public health crisis.[1] Obesity has been considered an important risk factor for the development of many chronic illnesses, including type 2 diabetes mellitus (T2DM).[2] As T2DM can lead to kidney damage, heart disease, stroke, and death, clinicians regularly suggest weight loss to prevent or reverse the onset of T2DM among individuals with obesity and overweight (BMI ≥25 kg/m^2).[3,4]

Although these guidelines apply across race and ethnicity, there is a particular interest in the role of obesity as a potential driver of T2DM prevalence and complications among minority populations.[5,6] At 49.6%, Black people have the highest rates of obesity in the United States.[7] Latine people follow with a 44.8% prevalence.[7] This is compared to a slightly lower 42.2% prevalence among White Americans.[7]

Black and Latine populations also have the highest rates of adult-onset T2DM. Critically however, their order in the prevalence hierarchy shifts. 13% percent of Black individuals have T2DM, compared with only eight percent of Whites, representing an increase of over 60%. The rate of T2DM among Latine individuals is a staggering 17%—more than 100% higher than that of Whites.[8,9]

Because the racial/ethnic group with the highest prevalence of obesity is not the one with the highest prevalence of T2DM, it is reasonable to suspect that the relationship between BMI and T2DM is inconsistent. Indeed, it may vary by race and/or ethnicity. Understanding the exact relationship between elevated BMI and T2DM prevalence by racial/ethnic group is essential for physicians aiming to provide accurate information about the preventive value of eradicating overweight and obesity among diverse populations.

☆ All data analyses were conducted at UCLA Office of Advanced Research Computing.

* Corresponding author. Tel: +949-229-5187.
E-mail address: sstrings@uci.edu (S. Strings).

https://doi.org/10.1016/j.puhe.2022.11.017

To date, we have only found one study dating from 1998 that examined the potential variations in the relationship between BMI and odds of diabetes by race. This study found, counterintuitively, that Black Americans had a higher risk of diabetes than White Americans at low BMIs. At higher body mass indices, they found no difference in risk of diabetes by race.[10] We have found no nationally representative studies examining the strength of the association between BMI and odds of T2DM across race/ethnicity that includes Latine populations. This is a critical oversight, given the burden of T2DM among Latine populations.

If elevations in BMI are driving the high rates of T2DM among Black and Latine populations, we would expect the strength of the association between obesity and T2DM to be strongest for Black, followed by Latine, populations. Therefore, in this study, we aimed to assess (1) the strength of the association between BMI and T2DM prevalence and (2) whether that relationship between BMI and T2DM risk varies by racial/ethnic category. This study may be a valuable tool in determining the relative risk for the development of T2DM that overweight and obesity pose for America's three largest racial/ethnic groups.

Research design and methods

Data were compiled from the continuous National Health and Nutrition Examination Survey (NHANES) from 1999–2000 through 2017–2018. NHANES contains health and nutrition data on a representative sample of non-institutionalized adults in the United States. Individuals aged ≥18 years were included. Data were analyzed for those that underwent the laboratory assessment (and thus had data on BMI and HbA_{1c}) and completed the demographic variable assessments ($N = 45{,}514$).

Weight and height were measured objectively by trained personnel. Weight was measured on a digital scale with participants in a standard examination gown. Height was measured to the nearest 1/10th of a centimeter. BMI was calculated using the standard equation of weight (kilogram)/height^2 (meter).

Race/ethnicity was coded as Black, White, or Latine (combining the Mexican American and Other Hispanic groups) from self-report. Gender and age (top-coded at 80 years) were also obtained via self-report.

For the outcome variable of odds of T2DM, we classified participants in each year into one of three risk tiers as recommended by the Centers for Disease Control and Prevention[10]: normal risk = HbA_{1c} <5.7%, prediabetes = HbA_{1c} 5.7%–6.4%, and T2DM = $HbA_{1c} \geq 6.5\%$.

The data violated ordered logistic regression assumptions, and therefore, generalized ordered logistic regression was used. The predictor variables were BMI, race/ethnicity, and a BMI × race/ethnicity interaction, plus covariates of gender, year, and age. As noted earlier, listwise deletion occurred for missing data. Analyses were conducted using Stata 17.0 (College Station, TX, USA). Following suggested NHANES procedures, sampling weights were adjusted for multiple years,[11] and stratification and clustering were incorporated as recommended.[12]

Results

Odds of prediabetes or T2DM compared with normal HbA_{1c} levels

Table 1 displays estimates of all models. Adjusting for age, gender, and year, each unit increase in BMI was associated with 10% higher odds of prediabetes or T2DM compared with having normal HbA_{1c} levels. There were race/ethnic differences in odds of prediabetes or T2DM. Compared with non-Hispanic White participants, non-Hispanic Black people and Latine participants had higher odds of prediabetes or T2DM.

Odds of T2DM compared with prediabetes and normal HbA_{1c} levels

Moreover, comparing to prediabetes or having normal HbA_{1c} levels, every unit increase in BMI was also associated with 11% higher odds of T2DM. Non-Hispanic Black and Latine participants had higher odds of T2DM than non-Hispanic White participants.

BMI × race/ethnicity interactions

As depicted in Table 1, the association of BMI with prediabetes or T2DM did not vary between non-Hispanic White and Latine participants. However, the association of BMI and (a) prediabetes or T2DM and (b) T2DM (Fig. 1) was weaker among non-Hispanic Black individuals compared with non-Hispanic White participants as signified by the significant interaction terms.

Discussion

In this nationally representative sample, Latine people had twice the odds of prediabetes or T2DM, and Black people had almost three times higher odds of prediabetes or T2DM compared with White people. BMI was also associated with higher odds of prediabetes or T2DM. However, the association between BMI and T2DM was altered substantially when race/ethnicity was considered. For White

Table 1
Associations between BMI and race/ethnicity (Model 1) and their interaction (Model 2) with T2DM, NHANES 1999–2018.

Predictor	Model 1		Model 2	
	Prediabetes and T2DM vs. normal glucose level OR (95% CI)	T2DM vs. prediabetes and normal glucose level OR (95% CI)	Prediabetes and T2DM vs. normal glucose level OR (95% CI)	T2DM vs. prediabetes and normal glucose level OR (95% CI)
BMI	1.10 (1.10–1.11)	1.11 (1.10–1.11)	1.11 (1.10–1.12)	1.12 (1.11–1.13)
Race/ethnicity				
Non-Hispanic White	1.00	1.00	1.00	1.00
Non-Hispanic Black	2.95 (2.71–3.20)	2.15 (1.93–2.39)	3.14 (2.88–3.42)	2.51 (2.24–2.83)
Latine	2.17 (1.98–2.38)	2.41 (2.13–2.72)	2.21 (2.02–2.43)	2.65 (2.30–3.04)
Body mass index × race/ethnicity				
Non-Hispanic White			1.00	1.00
Non-Hispanic Black			0.98 (0.97–0.99)	0.97 (0.95–0.98)
Latine			0.99 (0.98–1.00)	0.97 (0.96–0.99)
Age	1.07 (1.06–1.07)	1.06 (1.05–1.06)	1.07 (1.06–1.07)	1.06 (1.05–1.06)
Woman	0.75 (0.70–0.80)	0.62 (0.57–0.67)	0.76 (0.71–0.80)	0.62 (0.57–0.68)
Year	1.04 (1.03–1.05)	1.01 (1.01–1.02)	1.04 (1.03–1.05)	1.01 (1.01–1.02)

BMI, body mass index; CI, confidence interval; OR, odds ratio; T2DM, type 2 diabetes mellitus.

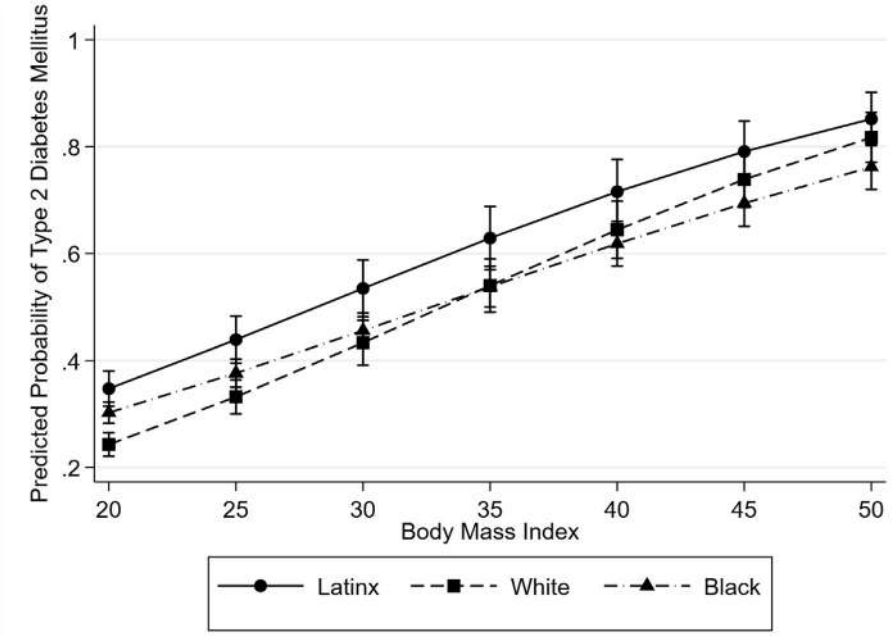


Fig. 1. Predicted probability of T2DM by BMI and race/ethnicity. Among all groups, there was a significant association between BMI and T2DM. Although non-Hispanic Black and Latine participants had higher predicted probability of T2DM vs. non-Hispanic White participants at similar BMI levels, the gradient between BMI and odds of T2DM was flatter among Black participants.

persons, there was a 10% greater odds of having prediabetes or T2DM with each BMI unit increase. This association was not statistically different for Latine persons. However, the odds of prediabetes/T2DM was significantly lower for Black people as a function of BMI. Moreover, the odds of T2DM with every unit increase in BMI were lower among Black and Latine people compared with White people.

Racial/ethnic differences in the associations between BMI and T2DM have not been fully examined. A study among women nurses found that the risk of T2DM incidence was higher among Black, Hispanic, and Asian women compared with White women at the same BMI.[13] Much of the discussion of this study focused on the larger than expected risk of T2DM among Asian participants, given similar BMI as White participants. However, increased risk of T2DM was also observed among other groups, suggesting that factors other than BMI are driving these higher odds.

These findings are an important contribution to our understanding of the relationship between BMI and T2DM. Existing scholarship has deemed minoritized people to be particularly at risk for T2DM, given their overrepresentation in the population of those with obesity. The current data, however, tell a different story. In fact, Black people appear to have *reduced* odds resulting from elevations in BMI; there is an 18% increased odds of T2DM between a Black person with a BMI in the "normal" category and a Black person with a BMI in the obesity category. This is lower than the equivalent for White persons, for whom obesity confers a 24% greater odds of T2DM than a person with a "normal" BMI.

We note the following strengths of this investigation. This study used nationally representative data and had clinically measured HbA_{1c} rather than self-reported data. The sample size was large, with 45,514 participants, and thus, any null results are likely not due to lack of power. There are also some limitations to this study. Given the constraints of NHANES, we chose to combine Mexican American and other Hispanic peoples under the umbrella of "Latine" for analytical parity. Among Latine populations, Mexican Americans have the highest rates of T2DM; thus, the relationships for Mexican Americans specifically may have been depressed. However, our goal was to identify differences among the major socially constructed and politically fused racial/ethnic groups in the United States, warranting a Latine category. Additional analyses are needed to understand intra-Latine variations in BMI-T2DM associations, as well as in other ethnicities. For example, there is very little research on the predictive utility of BMI for indigenous populations. What little exists suggests that BMI lacks explanatory power in these populations as well.[14] Second, this research did not assess other potential drivers of T2DM. Additional research is needed that considers factors such as stress, poverty, and food insecurity—all highly likely to be the upstream drivers of these associations and likely to vary by race/ethnicity.

This study has critical implications for preventive health in the context of T2DM. We argue that the wide variability in the association between BMI and T2DM across race serves as evidence that BMI has received unwarranted focus as a prime predictor of T2DM. Moreover, our findings directly contradict claims that the prevalence of obesity can explain differences in T2DM rates across racial and ethnic categories.

Although some may argue that these findings support the continued use of BMI as a screening tool among White populations, we would caution against this interpretation. White persons only comprise approximately 12% of the global population. Their dominant social and economic standing vis-à-vis people of color worldwide has meant that White people have among the lowest rates of T2DM and other chronic illnesses.[16] This power imbalance drives comparative analyses of health where White people are used as the referent. To the extent that the global public health community is invested in reducing racial/ethnic health inequities, there appears to be no need for a metric that realizes its best predictive ability among the world's privileged racial minority.

These results also have critical implications for research oriented toward improving health outcomes among Black Americans. BMI is not a good predictor of T2DM for Black individuals. Black people in

the study were nearly 115% more likely to have T2DM than White people, yet overweight and obesity were appreciably less influential to their rates of T2DM. These results suggest that the tremendous disparity in rates of T2DM among Black, Latine, and White Americans is not, in large part, attributable to differences in overweight and obesity. Indeed, BMI can be a faulty predictor of T2DM across all races by, for example, misclassifying muscle mass as fat.[15] Therefore, calls for weight loss as a corrective are likely misplaced.

Finally, the results reveal that BMI is not a fidelitous tool. Therefore, it may introduce bias and overestimation into interpretations of the factors propelling the prevalence of T2DM among minoritized groups and especially among Black people. Understanding the misjudgments and errors introduced by the continued utilization of BMI is crucial to eradicating health disparities. A growing body of research shows that the overreliance on BMI contributes to weight stigma, which can lead to worse health outcomes.[17,18] To the extent that people of color have higher BMIs on average than White persons, they are often targets of fatphobia. These realities are among the factors leading the WHO to release its 2017 report on the ills of weight stigma, as well as the NIH's 2022 webinar advising researchers on “moving beyond BMI” as a health metric.[19,20]

Instead of the continued overreliance on BMI, which creates a focus on the individual's need to change, researchers and clinicians should consider the social determinants of T2DM. Living below the poverty line; being food, employment, or housing insecure; or experiencing chronic stress have all been identified as upstream drivers of T2DM.[16,17] Racism may also be a factor contributing to elevations in diabetes prevalence.[18] Additional research is needed to better comprehend the factors contributing to the soaring rates of T2DM among Black individuals. Elevations in BMI alone cannot explain racial disparities in T2DM prevalence.

Author statements

Acknowledgments

Dr. Tomiyama’s time is supported by NIHgrants R01DK128575 and R01HL158555.

Ethical approval

Certified exempt under UCLA IRB#14-001686.

Funding

None declared.

Competing interests

None declared.

References

1. Office of the Surgeon General (US), Office of Disease Prevention and Health Promotion (US), Centers for Disease Control and Prevention (US), National Institutes of Health (US). *The surgeon general's call to action to prevent and decrease overweight and obesity*. 2001. Published online, https://www.ncbi.nlm.nih.gov/books/NBK44206/. [Accessed 6 April 2022].
2. Kahn SE, Hull RL, Utzschneider KM. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature* 2006;**444**(7121):840–6. https://doi.org/10.1038/nature05482.
3. National Institute of Diabetes and Digestive and Kidney Diseases. Managing diabetes. https://www.niddk.nih.gov/health-information/diabetes/overview/managing-diabetes.
4. Delahanty LM. Weight loss in the prevention and treatment of diabetes. *Prev Med* 2017;**104**:120–3. https://doi.org/10.1016/J.YPMED.2017.07.022.
5. Abraham PA, Kazman J ben, Zeno SA, Deuster PA. Obesity and african Americans: physiologic and behavioral pathways. *ISRN Obes* 2013;**2013**:1–8. https://doi.org/10.1155/2013/314295.
6. Obesity and African Americans - the Office of Minority Health. Accessed April 6, 2022. https://www.minorityhealth.hhs.gov/omh/browse.aspx?lvl=4&lvlid=25.
7. Adult obesity facts | overweight & obesity | CDC. Accessed April 6, 2022. https://www.cdc.gov/obesity/data/adult.html.
8. Centers for Disease Control and Prevention. Hispanic or Latino people and Type 2 diabetes.https://www.cdc.gov/diabetes/library/features/hispanic-diabetes.html#:~:text=Over%20their%20lifetime%2C%20US%20adults,it%20at%20a%20younger%20age.
9. U.S. Department of Health and Human Services Office of Minority Health. Diabetes and African Americans. https://minorityhealth.hhs.gov/omh/browse.aspx?lvl=4&lvlid=18.
10. Diabetes tests | CDC. Accessed April 6, 2022. https://www.cdc.gov/diabetes/basics/getting-tested.html.
11. NHANES tutorials - module 3 - weighting. Accessed April 6, 2022. https://wwwn.cdc.gov/nchs/nhanes/tutorials/module3.aspx.
12. NHANES Survey methods and analytic guidelines. Accessed April 6, 2022. https://wwwn.cdc.gov/nchs/nhanes/analyticguidelines.aspx.
13. Shai I, Jiang R, Manson JE, Stampfer MJ, Willett WC, Colditz GA, et al. Ethnicity, obesity, and risk of type 2 diabetes in women: a 20-year follow-up study. *Diabetes Care* 2006;**29**(7):1585–90.
14. Bambrick H. Relationships between BMI, waist circumference, hypertension and fasting glucose: rethinking risk factors in Indigenous diabetes. *Australian Indigenous Health Bulletin* 2005;**5**(4):1–15. Published online, http://www.healthinfonet.ecu.edu.au/uploads/resources/13783_bulletin_original_articles_bambrick.pdf. [Accessed 6 November 2022].
15. Rothman KJ. BMI-related errors in the measurement of obesity. *Int J Obes* 2008;**32**:56–9. https://doi.org/10.1038/ijo.2008.87.
16. Strings S, Ranchod YK, Laraia B, Nuru-Jeter A. Race and sex differences in the association between food insecurity and type 2 diabetes. *Ethn Dis* 2016;**26**(3):427–33. https://doi.org/10.18865/ED.26.3.427.
17. Hill J, Nielsen M, Fox MH. Understanding the social factors that contribute to diabetes: a means to informing health care and social policies for the chronically ill. *Perm J* 2013;**17**(2):67–72. https://doi.org/10.7812/TPP/12-099.
18. Golden SH, Joseph JJ, Hill-Briggs F. Casting a health equity lens on endocrinology and diabetes. *J Clin Endocrinol Metab* 2021;**106**(4):e1909–16. https://doi.org/10.1210/CLINEM/DGAA938.

Public Health 215 (2023) 106–117

Contents lists available at ScienceDirect

Public Health

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

Review Paper

The effectiveness of prenatal and postnatal home visits by paramedical professionals and women's group meetings in improving maternal and child health outcomes in low and middle-income countries: a systematic review and meta-analysis

M. Gupta [a, *], A. Bansal [a], V. Chakrapani [b], N. Jaiswal [a], T. Kiran [a]

[a] Postgraduate Institute of Medical Education and Research, Chandigarh, India
[b] Centre for Sexuality and Health Research and Policy (C-SHaRP), Chennai, India

ARTICLE INFO

Article history:
Received 16 May 2022
Received in revised form
22 November 2022
Accepted 29 November 2022
Available online 20 January 2023

Keywords:
Effectiveness
Home visits
Meta-analysis
Prenatal
Postnatal
Paramedical professionals
Systematic review
Women group meetings

ABSTRACT

Objective: To assess the effectiveness of prenatal and postnatal home visits (HVs) and women group meetings (WGMs) by paramedical professionals to improve maternal and child health outcomes in low- and middle-income countries (LMICs).
Study design: Systematic review and meta-analysis.
Methods: We conducted a systematic review of trials published till December 2020, as per registered protocol in The International Prospective Register of Systematic Reviews (PROSPERO) (CRD42018091968). Outcomes were neonatal mortality rate (NMR), maternal mortality ratio (MMR), the incidence of low birth weight, and still birth rate (SBR). The Cochrane Pregnancy and Childbirth Group's Trials Register, Cochrane Central Register of Controlled Trials, PubMed, and Excerpta Medica Database (EMBASE) were searched. Pooled results were estimated using random-effects meta-analysis in RevMan version 5.2.
Results: Twenty-five trials met the inclusion criteria. HVs were the key intervention in 12, WGMs in 11, and both interventions in 2 trials. The pooled estimates have shown that NMR was significantly reduced by HVs (OR 0.77, confidence interval [CI]: 0.67–0.90, $P = 0.0007$, $I^2 = 77\%$) and WGMs (OR 0.76, CI: 0.65–0.90, $P = 0.001$, $I^2 = 71\%$). SBR was significantly reduced by HVs (OR 0.77, CI: 0.70–0.85; $P < 0.001$, $I^2 = 0\%$). Subgroup analysis of studies in which more than 10% of pregnant women participated in the WGMs showed significant reduction in NMR (OR 0.67, CI 0.58–0.77, $P = 0.00001$, $I^2 = 31\%$) and MMR (OR 0.55, CI 0.36–0.84, $P = 0.005$, $I^2 = 27\%$). Two studies reported improvement in birth weight by HVs.
Conclusions: HVs and WGMs (with >10% pregnant women) by paramedical professionals are effective strategies in reducing the NMR and MMR in LMICs. HVs were also effective in reducing SBR.

Introduction

The sustainable development goals (SDGs) aim to reduce the under-five mortality rate to 25 deaths per 1000 live births across the world by the year 2030.[1] According to the World Health Organization (WHO), nearly 6 million children under the age of 5 years died in 2016.[2] To achieve the SDGs, there is a need to reduce the neonatal mortality rate (NMR) as it constitutes 41% of the child mortality, globally.[3] The low- and middle-income countries (LMICs) contribute to about three-fourths of the global NMR.[4] Major causes of neonatal mortality in low resource settings are birth asphyxia, birth injuries, preterm births, low birth weight, neonatal sepsis, and hypothermia.[5] Child mortality is associated with maternal well-being. Providing skilled care to women during, prenatal, natal, and postnatal periods, greatly contributes to child survival. The revised targets for maternal mortality ratio (MMR) as part of SDGs are to reduce MMR to less than 70 per 100,000 live births.[6] Globally, more than 800 women die every day during labor, or within the first 24 h after delivery, with nearly 99% of such deaths occurring in the LMICs.[6]

Several studies have examined the role of home visits (HVs) by paramedical professionals (community health worker [CHW], or community volunteer [CV], or auxiliary nurse midwives [ANM] or multipurpose workers [MPW] or accredited social health activists

* Corresponding author. Community Medicine, Department of Community Medicine and School of Public Health, Postgraduate Institute of Medical Education and Research, Sector 12, Chandigarh, 160012, India. Tel.: +91 7087008223.
E-mail address: madhugupta21@gmail.com (M. Gupta).

https://doi.org/10.1016/j.puhe.2022.11.023
0033-3506/

(ASHA) or nurse) during prenatal and postnatal periods in improving the birth outcomes, including NMR and MMR, in LMICs including Pakistan, Uganda, and Ethiopia.[6–9] India strengthened its community health services through implementation of National Rural Health Mission in 2005, and hiring approximately 800,000 ASHAs at village level, making it a largest CHW program in the world.[10] The lady health worker program in Pakistan provides HVs in the community to support immunization, nutrition, and family planning services. The CHWs in India provide 3–4 visits during the antenatal period of the pregnant women and 3–4 visits during the postnatal period. During the antenatal period, women are provided education regarding healthy diet during pregnancy, birth preparedness, and institutional delivery. During postnatal care, counseling includes self-hygiene and hygiene of the baby, exclusive breastfeeding, special care requirements of babies with low birth weight, and immunization of the child till 6 weeks after the delivery.[9] These efforts indicate that LMICs have done lot of efforts in improving the MCH outcomes through paramedical professionals. On the other hand, such community health programs have put significant burden on the financial resources of the LMICs. One study has reported the annual cost of the kit provided to a CHW for visiting pregnant women per 100,000 population ranges from 15 USD to 116 USD in LMICs.[11] In view of this, the effectiveness of HVs by paramedical professionals to improve MCH outcomes need to be reviewed and consolidated. Hence, a systematic review was conducted to assess the current evidence on the role of prenatal and postnatal HVs by paramedical professionals in reducing NMR, MMR, and incidence of low birth weight (LBW), in LMICs. This is the most updated review on this topic in comparison with other systematic reviews.[12–14]

Methods

This systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines[15] (Supplementary Table 1). The review protocol was registered in the PROSPERO international prospective register of systematic reviews (CRD42018091968).[16]

Type of studies

Cochrane search strategy for randomized controlled trials (RCTs) or quasi-experimental studies on HVs by paramedical professionals or the WGMs focusing on pregnant women conducted in LMICs, and published till December, 2020 in Cochrane Pregnancy and Childbirth Group's Trials Register, Cochrane Central Register of Controlled Trials, PubMed, and EMBASE was used.[17] No language restrictions were done during the searches. Studies were eligible for inclusion in the review if: a) the assignment of study participants to both intervention and control groups was either randomized or chosen purposively, b) the study intervention had its major focus on the effectiveness of HVs or WGMs by paramedical professionals on pregnancy outcomes including NMR, MMR, and incidence of LBW, c) there was at least one outcome measure evaluated in the intervention and control group, and d) the data were either reported in a useable form (Odds Ratio/Risk Ratio [OR/RR] or numbers from which these measures could be generated) or useable data could be retrieved from the papers.

Exclusion criteria

- The studies which were conducted in high-income countries.[18]
- Studies with a focus on interventions/outcomes other than those mentioned in the protocol of this review.[16]
- Studies other than randomized control trials/Quasi-experimental.

Study participants

The study population was antenatal or postnatal women who had received HVs by any paramedical professional (CHW, or CV, or ANM or MPW or ASHA or nurse), and women in their reproductive age group who had participated in WGMs focusing on issues related to care during pregnancy and/or postnatal period, with any number of frequencies. The control group included the group of women who did not receive any of the above-stated interventions and had received routine maternal and child care in the national health programs.

Type of interventions

The interventions refer to the counseling or health education provided during the HVs in prenatal period as well as during WGMs as mentioned in the registered protocol.[16] These include providing information or counseling on one or more of the following: (a) healthy diet and cooking practices during pregnancy, (b) personal hygiene, sleeping hours during pregnancy, (c) regular prenatal check-ups and follow up in health facility, (d) birth preparedness, and social support during delivery, (e) creating awareness of danger signs during pregnancy, delivery and neonatal period, and offering early referral for management of complications, and (f) creating awareness on schemes/benefits by the government for institutional delivery or any other related issues.

Outcome measures

The primary outcomes considered in this review were (a) NMR, (b) MMR, and (c) incidence of LBW. The secondary outcomes were (a) still birth rate (SBR) and (b) institutional delivery rate (IDR).

Information sources

The search strategy was prepared by two researchers. The reference lists of all identified articles on interventions were checked to identify relevant studies. Also, citations tracking of prominent researchers working in the field of HVs/WGMs and pregnancy outcomes was conducted to identify relevant articles. Further, hand-searching of the contents of reputed obstetric/public health journals and conference proceedings was also conducted. Relevant articles and reports were searched in Google, Google Scholar, and in databases of agencies such as UNICEF and WHO. Only those trials conducted in LMICs were included.[18]

The medical subject headings words used for searches are given in Table 1.

Study selection

Two phase screening of the studies was done. In the first phase two reviewers (AB, NJ) independently screened the studies according to their titles and abstracts, after removing the duplications using Zotero software version 4.0.[19] In the second phase, full text review of the shortlisted studies was done by two reviewers (AB, NJ). Differences in the opinion regarding inclusion or exclusion of a study were resolved by appealing to the third and fourth reviewers (MG, VC) in both the phases. The references list of each selected paper was cross-checked to identify additional studies that met the selection criteria.

Table 1
Search strategy, systematic review, impact of home visits, or group meetings by paramedical staff or community-based workers during pregnancy to improve maternal and child health outcomes.

Low and middle income countries	('low-income economies' OR 'lower middle-income economies' OR 'middle income economies' OR 'developing countries' [MeSH Terms] OR ('developing' [All Fields] AND 'countries' [All Fields]) OR 'developing countries' [All Fields]) OR ('developing countries' [MeSH Terms] OR ('developing' [All Fields] AND 'countries' [All Fields]) OR 'developing countries' [All Fields] OR ('developing' [All Fields] AND 'country' [All Fields]) OR 'developing country' [All Fields]) OR ('developing countries' [MeSH Terms] OR ('developing' [All Fields] AND 'countries' [All Fields]) OR 'developing countries' [All Fields] OR ('underdeveloped' [All Fields] AND 'countries' [All Fields]) OR 'underdeveloped countries' [All Fields]) OR ('developing countries' [MeSH Terms] OR ('developing' [All Fields] AND 'countries' [All Fields]) OR 'developing countries' [All Fields] OR ('underdeveloped' [All Fields] AND 'country' [All Fields]) OR 'underdeveloped country' [All Fields]) OR (emergent [All Fields] AND Countries [All Fields]) OR (emergent [All Fields] AND country [All Fields]) OR ('developing countries' [MeSH Terms] OR ('developing' [All Fields] AND 'countries' [All Fields]) OR 'developing countries' [All Fields] OR ('developing 'nation' [All Fields]) OR 'developing nation' [All Fields]) OR (underdeveloped [All Fields] AND 'nation' [All Fields]))
Search by name of countries	(Afghanistan or Albania or Algeria or Angola or Antigua or Barbuda or Argentina or Armenia or Armenian or Aruba or Azerbaijan or Bahrain or Bangladesh or Barbados or Benin or Byelarus or Byelorussian or Belarus or Belorussian or Belorussia or Belize or Bhutan or Bolivia or Bosnia or Herzegovina or Hercegovina or Botswana or Brazil or Brasil or Bulgaria or 'Burkina Faso' or 'Burkina Fasso' or 'Upper Volta' or Burundi or Urundi or Cambodia or 'Khmer Republic' or Kampuchea or Cameroon or Cameroons or Cameron or Camerons or 'Cape Verde' or 'Central African Republic' or Chad or Chile or China or Colombia or Comoros or 'Comoro Islands' or Comores or Mayotte or Congo or Zaire or 'Costa Rica' or 'Cote d'Ivoire' or 'Ivory Coast' or Croatia or Cuba or Cyprus or Czechoslovakia or 'Czech Republic' or Slovakia or 'Slovak Republic' or Djibouti or 'French Somaliland' or Dominica or 'Dominican Republic' or 'East Timor' or 'East Timur' or 'Timor Leste' or Ecuador or Egypt or 'United Arab Republic' or 'El Salvador' or Eritrea or Estonia or Ethiopia or Fiji or Gabon or 'Gabonese Republic' or Gambia or Gaza or Georgia or Georgian or Ghana or 'Gold Coast' or Greece or Grenada or Guatemala or Guinea or Guam or Guiana or Guyana or Haiti or Honduras or Hungary or India or Maldives or Indonesia or Iran or Iraq or 'Isle of Man' or Jamaica or Jordan or Kazakhstan or Kazakh or Kenya or Kiribati or Korea or Kosovo or Kyrgyzstan or Kirghizia or 'Kyrgyz Republic' or Kirghiz or Kirgizstan or 'Lao PDR' or Laos or Latvia or Lebanon or Lesotho or Basutoland or Liberia or Libya or Lithuania or Macedonia or Madagascar or 'Malagasy Republic' or Malaysia or Malaya or Malay or Sabah or Sarawak or Malawi or Nyasaland or Mali or Malta or 'Marshall Islands' or Mauritania or Mauritius or 'Agalega Islands' or Mexico or Micronesia or 'Middle East' or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Morocco or Ifni or Mozambique or Myanmar or Myanma or Burma or Namibia or Nepal or 'Netherlands Antilles' or 'New Caledonia' or Nicaragua or Niger or Nigeria or 'Northern Mariana Islands' or Oman or Muscat or Pakistan or Palau or Palestine or Panama or Paraguay or Peru or Philippines or Philipines or Phillipines or Phillippines or Poland or Portugal or 'Puerto Rico' or Romania or Rumania or Roumania or Russia or Russian or Rwanda or Ruanda or 'Saint Kitts' or 'St Kitts' or Nevis or 'Saint Lucia' or 'St Lucia' or 'Saint Vincent' or 'St Vincent' or Grenadines or Samoa or 'Samoan Islands' or 'Navigator Island' or 'Navigator Islands' or 'Sao Tome' or 'Saudi Arabia' or Senegal or Serbia or Montenegro or Seychelles or 'Sierra Leone' or Slovenia or 'Sri Lanka' or Ceylon or 'Solomon Islands' or Somalia or Sudan or Suriname or Surinam or Swaziland or Syria or Tajikistan or Tadzhikistan or Tadjikistan or Tadzhik or Tanzania or Thailand or Togo or 'Togolese Republic' or Tonga or Trinidad or Tobago or Tunisia or Turkey or Turkmenistan or Turkmen or Uganda or Ukraine or Uruguay or USSR or 'Soviet Union' or 'Union of Soviet Socialist Republics' or Uzbekistan or Uzbek or Vanuatu or 'New Hebrides' or Venezuela or Vietnam or 'Viet Nam' or 'West Bank' or Yemen or Yugoslavia or Zambia or Zimbabwe or Rhodesia)
Home visits	'house calls'[MeSH Terms] OR ('house'[All Fields] AND 'calls'[All Fields]) OR 'house calls'[All Fields] OR ('home'[All Fields] AND 'visits'[All Fields]) OR 'home visitation'[All Fields]
Paramedical professionals	'Paramedical professionals' [MeSH Terms] OR 'Paramedical [All Fields]AND professionals [All Fields]' OR 'staff nurses [MeSH Terms]' OR 'staff [AllFields]ANDnurses[All Fields]"communityhealthworkers' [MeSH Terms] OR 'health workers' [All Fields]
Women groups	'women groups'[MeSH Terms] OR ('women help groups'[All Fields] AND 'female groups'[All Fields]) OR 'group meetings'[All Fields]
Pregnant women	pregnant women'[MeSH Terms] OR ('pregnant'[All Fields] AND 'women'[AllFields])OR'pregnantwomen'[AllFields] AND ('parturition'[MeSH Terms] OR 'parturition'[All Fields] OR 'birth'[AllFields]) AND outcomes [All Fields]
Pregnancy outcome	'pregnancy outcome'[MeSH Terms] OR ('pregnancy'[All Fields] AND 'outcome'[All Fields]) OR 'pregnancy outcome'[All Fields]
low birth weight	'infant, low birth weight' [MeSH Terms] OR ('infant'[All Fields] AND 'low'[All Fields] AND 'birth'[All Fields] AND 'weight'[All Fields]) OR 'low birth weight infant'[All Fields] OR ('low'[All Fields] AND 'birth'[All Fields] AND 'weight'[All Fields]) OR 'low birth weight'[All Fields]
Neonatal mortality	'infant mortality'[MeSH Terms] OR ('infant'[All Fields] AND 'mortality'[All Fields]) OR 'infant mortality'[All Fields] OR ('neonatal'[All Fields] AND 'mortality'[All Fields]) OR 'neonatal mortality'[All Fields]
Maternal Mortality	'maternal mortality'[MeSH Terms] OR ('maternal'[All Fields] AND 'mortality'[All Fields]) OR 'maternal mortality'[All Fields]
Randomized Control trials	'random allocation'[MeSH Terms] OR ('random'[All Fields] AND 'allocation'[All Fields]) OR 'random allocation'[All Fields] OR 'randomized'[All Fields]) AND ('prevention and control'[Subheading] OR ('prevention'[All Fields] AND 'control'[All Fields]) OR 'prevention and control'[All Fields] OR 'control'[All Fields] OR 'control groups'[MeSH Terms] OR ('control'[All Fields] AND 'groups'[All Fields]) OR 'control groups'[All Fields]) AND ('Trials'[Journal] OR 'trials'[All Fields]

Data collection

Data were extracted from the selected studies by two independent reviewers (AB, NJ) using standardized data extraction form. Discrepancies in the filled-in data extraction form were resolved by discussing with third and fourth reviewer (MG, VC). From each eligible study, we extracted relevant information on the study design, participant settings, methods, interventions, and outcomes.

Assessment of risk of bias

The studies were classified as having 'Low risk', 'Unclear' or 'High risk' of bias using the Cochrane Collaboration's tool for assessing the risk of bias in randomized control trials by the review authors (AB, MG).[20] Funnel plots were created to depict the risk of publication bias.

Data analysis

The data of all the eligible trials were divided according to the intervention provided to the participants (HVs or WGMs). The results were pooled for meta-analysis using RevMan software version 5.2[21] and analyzed separately for both the interventions. The individually randomized trials in which the sample size or the event numbers were not distinctly mentioned were not pooled; however, the results of such studies were discussed in the review. We used Der Simonian and Laird random-effects models to develop mean summary estimates and 95% confidence intervals (CIs)

pooled estimates.[22] Forest plot was used to present the pooled estimates for the primary outcomes where the studies were found in sufficient numbers. Substantial heterogeneity (I^2 < 60%) was investigated using subgroup analysis wherever we could find population level predictors (like proportion of pregnant women attending the WGMs) which could impact the outcomes. The certainty of evidence generated by the included studies was assessed by GRADE approach using GRADEpro application.[23] The certainty of evidence was categorized as 'high', 'moderate', or 'low'. The evidence was downgraded where the heterogeneity was found to be more than 60%.

Results

We identified 2011 research papers in the first stage. PRISMA flow diagram showing studies' selection process is given in Fig. 1. After removing the duplicated (N = 328) and ineligible studies (N = 1665) and inclusion of cross reference studies (N = 7), 25 studies were included in this review. Six studies were conducted in India,[24–29] four each in Bangladesh[30–33] and Pakistan,[34–37] three in Tanzania,[38–40] two each in South Africa,[41,42] Nepal,[43,44] and Malawi,[45,46] and one each in China[47] and Ghana.[48] There were 23 cluster randomized control trials,[24,26–36,38–48] one individual randomized control trial,[25] and one quasi-experimental trial.[37] Details of the selected studies is given in Supplementary Table 2. Overall, more than 250,000 women either pregnant or in the reproductive age group were the participants in the selected trials.

Interventions

HVs by paramedical professionals was the key intervention in 12 studies conducted in Pakistan, Bangladesh, Ghana, South Africa, India, Tanzania, and Malawi.[24–26,30,34,35,38–42,48] Participatory learning and action approaches by conducting WGMs with women in the reproductive age group in the villages by the female facilitators were the key intervention in 11 studies.[27–29,32,33,36,37,43,44,46,47] These studies were conducted in Bangladesh, India, Nepal, Pakistan and Malawi and China.[27,29,31–34,37,43,45–47] Two trials had used HVs as intervention in one arm and WGMs in other arm and analyzed them against a common control.[31,45]

Information education communication activities like group education sessions for awareness generation by involving media were reported to be conducted along with HVs in studies conducted in India, Malawi, and in a multicountry study (India, Pakistan, Kenya, Zambia, and Argentina)[25,26,31,45] and also with trials from Nepal and Bangladesh that used WGMs.[27,43,47] Four studies with HVs as the intervention had also focused on strengthening of the health systems by training health care staff (birth attendants) for better provision of healthcare[26,35,47,48] (Supplementary Table 2).

Outcomes

Primary outcomes

NMR was the primary outcome in nineteen out of the 25 studies.[24,26–37,39,43,45–48] Two studies reported MMR[45,47] and three studies reported incidence of LBW as the primary outcome[25,41,42] (Supplementary Table 2).

Secondary outcomes

Ten studies had MMR as the secondary outcome.[27–29,32–35,37,43,46] SBR was the secondary outcome in thirteen studies,[24,26–30,32–35,43,46,47] IDR in sixteen studies[27–30,32,33,35–40,43–45,48] and behavior changes related to neonatal health in five studies[27,32,38,41,42] (Supplementary Table 2).

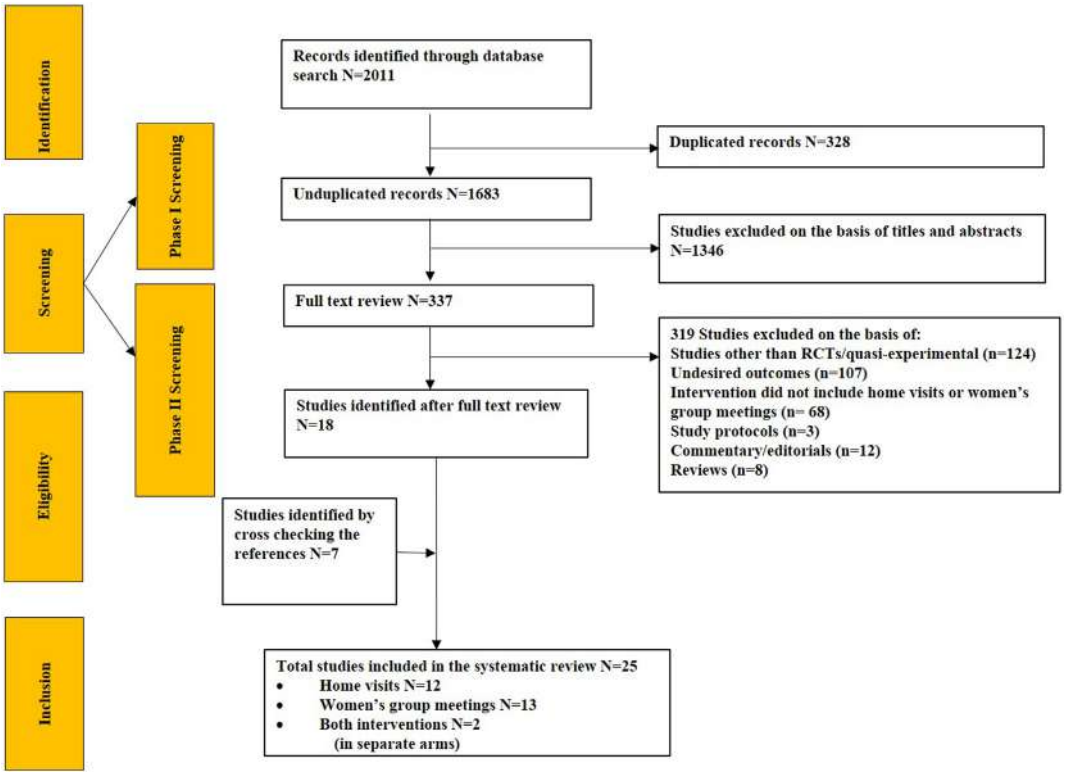


Fig. 1. PRISMA flow diagram of studies on home visits or women group meetings by paramedical professionals to improve birth outcomes in low-and-middle-income countries. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Table 2
Pooled results of various birth outcomes.

Study	Number of participants	Events in intervention group	Events in control group	Rate in intervention group	Rate in control group	Odds ratio [CI 95%]	Standard Error	Weight (%)	Pooled OR [95% CI]	*P* value
A. Studies with home visits as an intervention										
1. Neonatal mortality rate (NMR)	Live births (N)	Neonatal deaths (N)	Neonatal deaths (N)	NMR (per thousand live births)	NMR (per thousand live births)				0.77 [0.67, 0.90]	0.0007
1. Baqui et al., 2008[a,30]	30,119	561	696	29.2	43.2	0.66 (0.47–0.93)	0.1732	8.7%		
2. Bhutta et al., 2008[33]	5542	121	156	41.3	59.8	0.72 (0.56–0.91)	0.1242	11.2%		
3. Bhutta et al., 2011[34]	23,033	517	540	43	49.1	0.85 (0.76–0.96)	0.0571	14.8%		
4. Dramstadt et al., 2010[30]	9857	111	146	24	27.9	0.87 (0.68–1.12)	0.1257	11.2%		
5. Hanson et al., 2015[38]	47,688	749	679	31	30	1.0 (0.9–1.2)	0.0538	15.0%		
6. Kirkwood et al., 2013[47]	15,619	230	252	29.8	31.9	0.92 (0.75–1.12)	0.1042	12.4%		
7. Kumar et al., 2008[23]	2601	64	91	41	84.2	0.46 (0.35–0.60)	0.1394	10.4%		
8. Lewycka et al., 2013[a,45]	9570	95	147	NR	NR	0.78 (0.48–1.27)	0.2477	5.9%		
9. Rasaily et al., 2020[25]	12,322	173	194	26.5	33.5	0.75 (0.57–0.99)	0.14	10.4%		
2. Maternal mortality ratio (MMR)	Pregnant women (N)	Deaths (N)	Deaths (N)	MMR (per lakh live births)	MMR (per lakh live births)				0.76 [0.57, 1.02]	0.07
1. Bhutta et al., 2008[33]	5542	5	8	NR	NR	0.56 [0.18, 1.70]	0.5707	6.8%		
2. Bhutta et al., 2011[34]	23,033	71	78	NR	NR	0.83 [0.60, 1.15]	0.1646	82.0%		
3. Lewycka et al., 2013[a,45]	9681	18	29	NR	NR	0.48 [0.20, 1.15]	0.4467	11.1%		
3. Stillbirths	Births (N)	Stillbirths (N)	Stillbirths (N)	Stillbirth rate (per thousand live births)	Stillbirth rate (per thousand live births)				0.77 [0.70, 0.85]	<0.00001
1. Bhutta et al., 2008[33]	5842	132	168	43.1	60.5	0.66 (0.53–0.83)	0.1119	21.6%		
2. Bhutta et al., 2011[34]	24,085	489	563	39.1	48.7	0.79 (0.68–0.92)	0.0765	46.3%		
3. Dramstadt et al., 2010[30]	9857	113	109	NR	NR	0.91 [0.70, 1.19]	0.1358	14.7%		
4. Kumar et al., 2008[23]	2724	59	64	39.1	54.1	0.72 (0.52–1.00)	0.166	9.8%		
5. Rasaily et al., 2020[25]	12,521	98	101	14.8	17.3	0.81 (0.56–1.16)	0.1883	7.6%		
4. Institutional delivery	Pregnant women (N)	Deliveries (N)	Deliveries (N)	Rate						
%	Rate									
%				1.20 [0.98, 1.47]	0.08					
1. Bhutta et al., 2011[34]	4474	1272	936	54%	44%	1.53 [1.36, 1.72]	0.0602	18.0%		
2. Dramstadt et al., 2010[30]	4103	350	287	20.20%	12.10%	1.31 [0.64, 2.68]	0.3655	5.6%		
3. Geldsetzar et al., 2019[40]	1374	NR	NR	96.10%	92.70%	0.54 [0.30, 0.97]	0.2999	7.3%		
4. Hanson et al., 2015[38]	15,573	6412	5665	82%	75%	1.50 [1.20, 1.88]	0.1139	15.5%		
5. Kirkwood et al., 2013[47]	15,980	5373	5539	NR	NR	0.97 [0.81, 1.16]	0.092	16.6%		
6. Kumar et al., 2008[23]	NR	NR	NR	19.7	14	1.41 [0.93, 2.14]	0.2123	10.6%		
7. Lewycka et al., 2013[a,45]	1567	400	284	NR	NR	1.04 [0.86, 1.26]	0.097	16.4%		
8. Penfold et al., 2014[37]	510	187	166	73%	65%	1.40 [0.90, 2.18]	0.2254	10.0%		
B. Studies with women's group meetings as the intervention										
1. Neonatal mortality rate (NMR)	Live births (N)	Neonatal deaths (N)	Neonatal deaths (N)	NMR (per thousand live births)	NMR (per thousand live births)				0.76 [0.65, 0.90]	0.001
1. Azad et al., 2010[31]	29,889	515	557	NR	NR	0.90 [0.73, 1.11]	0.1068	12.6%		
2. Baqui et al., 2008[a,30]	31,675	807	696	45.2	43.5	0.95 [0.69, 1.31]	0.1632	10.0%		
3. Colbourn et al., 2013[26]	19,986	10,055	9931	31	28.4	0.90 [0.75, 1.08]	0.093	13.3%		
4. Fottrell et al., 2013[32]	17,421	18	271	21.3	30.1	0.62 [0.43, 0.89]	0.1867	8.9%		
5. Lewycka et al., 2013[a,45]	9640	110	147	23.5	29.6	0.59 [0.36, 0.97]	0.2521	6.6%		
6. Manandhar et al., 2004[42]	6125	76	119	26.2	36.9	0.70 [0.53, 0.92]	0.1419	10.9%		
7. Midhet et al., 2010[43]	1635	24	43	32.4	48	0.45 [0.27, 0.74]	0.2598	6.3%		
8. More et al., 2012[28]	15,703	132	88	16.62	11.34	1.42 [0.99, 2.04]	0.184	9.0%		
9. Tripathy et al., 2010[26]	18,449	406	531	42.9	59.1	0.68 [0.59, 0.78]	0.0724	14.2%		
10. Tripathy et al., 2016[27]	7042	108	151	30	43.9	0.54 [0.36, 0.81]	0.2069	8.1%		
2. Maternal mortality ratio (MMR)	Pregnant women (N)	Deaths (N)	Deaths (N)	MMR (per lakh live births)	MMR (per lakh live births)				0.68 [0.47, 1.00]	0.05
1. Azad et al., 2010[31]	29,889	55	32	NR	NR	1.74 (0.97–3.13)	0.2981	14.3%		
2. Colbourn et al., 2013[45]	19,986	25	22	251.7	218.8	0.91 (0.51–1.63)	0.2954	14.4%		
3. Fottrell et al., 2013[32]	17,421	14	23	153.4	276.1	0.74 (0.34–1.64)	0.3968	11.3%		
4. Lewycka et al., 2013[a,45]	9640	14	29	299	585	0.26 (0.10–0.68)	0.4875	9.1%		
5. Manandhar et al., 2004[42]	6125	2	11	69	341	0.22 (0.05–0.90)	0.7559	5.0%		
6. More et al., 2012[28]	15,703	20	24	NR	NR	0.81 (0.45–1.47)	0.3032	14.1%		
7. Osrin et al., 2003[43]	NR	NR	NR	NR	NR	0.22 [0.05, 0.97]	0.7559	5.0%		
8. Tripathy et al., 2010[26]	18,449	49	60	517.5	668.1	0.70 (0.46–1.07)	0.2142	17.1%		
9. Tripathy et al., 2016[27]	7042	8	12	222	348.9	0.63 (0.25–1.42)	0.4716	9.5%		
3. Stillbirths	Births (N)	Stillbirths (N)	Stillbirths (N)	Stillbirth rate (per thousand live births)	Stillbirth rate (per thousand live births)				0.96 [0.84, 1.09]	0.53
1. Azad et al., 2010[31]	29,889	542	521	NR	NR	1.00 (0.82–1.21)	0.1013	16.3%		
2. Colbourn et al., 2013[45]	20,576	316	274	30.8	26.5	0.81 (0.65–1.02)	0.1123	14.1%		
3. Fottrell et al., 2013[32]	17,940	287	232	31.6	23.5	1.07 (0.75–1.53)	0.1813	6.6%		
4. Manandhar et al., 2004[42]	6275	73	77	24.6	23.3	1.06 (0.76–1.47)	0.1698	7.3%		

Table 2 (*continued*)

Study	Number of participants	Events in intervention group	Events in control group	Rate in intervention group	Rate in control group	Odds ratio [CI 95%]	Standard Error	Weight (%)	Pooled OR [95% CI]	*P* value
5. More et al., 2012[28]	18,197	73	85	7.97	9.4	0.66 (0.46–0.93)	0.1842	6.4%		
6. Tripathy et al., 2016[27]	7219	97	80	26.2	22.7	1.27 (0.80–2.01)	0.1018	16.2%		
7. Tripathy et al., 2010[26]	18,775	301	280	30.8	30.2	1.05 (0.86–1.28)	0.2358	4.1%		
4. Institutional deliveries	Pregnant women (N)	Deliveries (N)	Deliveries (N)	Rate%	Rate%				1.20 [0.98, 1.47]	0.07
1. Azad et al., 2010[31]	29,889	226	302	NR	NR	0.97 (0.77, 1.24)	0.1178	17.5%		
2. Fottrell et al., 2013[32]	17,940	NR	NR	26.8	27.7	1.05 (0.88–1.25)	0.0901	19.3%		
3. Lewycka et al., 2013[a,45]	19,123	4733	4891	NR	NR	1.27 (0.95–1.71)	0.1481	15.4%		
4. Manandhar et al., 2004[42]	6275	201	66	NR	NR	3.55 (1.56–8.05)	0.4195	4.8%		
5. Midhet et al., 2010[35]	1858	NR	NR	NR	NR	1.3 (0.7, 2.5)	0.3158	7.3%		
6. More et al., 2012[28]	15,192	6602	6573	NR	NR	0.92 (0.58–1.47)	0.2354	10.4%		
7. Osrin et al., 2003[43]	NR	NR	NR	NR	NR	3.54 (1.56–8.03)	0.4181	4.8%		
8. Tripathy et al., 2010[26]	18,335	1364	1811	NR	NR	0.89 (0.51–1.53)	0.2841	8.4%		
9. Tripathy et al., 2016[27]	7219	2364	1943	NR	NR	1.20 (0.81–1.78)	0.2005	12.2%		

[a] Studies with both the interventions (home visits and women's group) and analyzed separately against the control arm/cluster.

Summary of the findings

The studies where RR or OR was reported or could be calculated were pooled, and forest plot was generated for the primary outcomes. As there were only two relevant studies on LBW incidence (an outcome of interest for this review),[41,42] these were not pooled. Pooled estimates of the studies (if at least three similar studies were found), with NMR, MMR, incidence of LBW, SBR, and IDR as the outcomes is presented in Table 2.

The outcomes of the systematic review are summarized below:

NMR

Of the 19 trials (18 cluster RCTs and 1 quasi-experimental) that have reported NMR as a primary outcome, 13 trials were conducted in the South East Asian Region,[24,26–36,43] four in the African continent[39,45,46,48] and one was a multicountry study.[47] Seven trials had only HVs, eight trials only WGMs and two had both the interventions implemented in different arms. Pooling of nine trials with HVs as the intervention (OR 0.77 [95% CI 0.67, 0.90]; $P = 0.0007$, $I^2 = 77\%$)[24,26,30,31,34,35,39,45,48]; and ten trials with WGMs as the intervention (OR 0.76 [95% CI 0.65, 0.90]; $P = 0.001$, $I^2 = 71\%$)[27–29,31–33,36,43,45,46] showed a statistically significant decline in NMR (Table 2; Fig. 2a and c).

A study by Pasha et al.[47] could not be pooled because of unavailability of the data needed. It reported a non-significant effect of WGMs on reducing NMR ($P = 0.33$). A quasi-experimental study by Memon et al.[37] conducted in Pakistan reported a significant reduction in NMR (OR 0.58; 95% CI: 0.48–0.68; $P = 0.03$) in the intervention group having both HVs and WGMs as an intervention.

MMR

Of the 25 studies, twelve had MMR as the outcome (two as primary and ten as a secondary outcome).[27–29,32–35,43–47] Two trials had HVs, eight trials WGMs and one trial had both the interventions in different study arms. Pooling of three trials with HVs as an intervention (OR 0.76 [95% CI 0.57, 1.02]; $P = 0.07$,

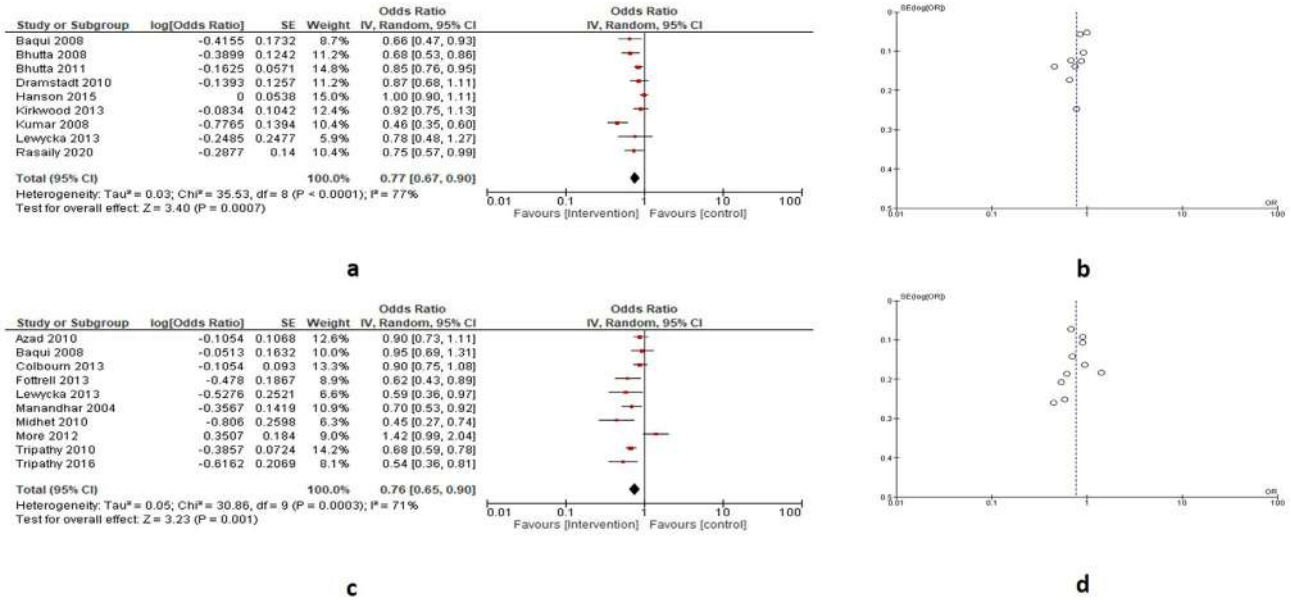


Fig. 2. Meta-analysis of effect of home visits and women's group intervention by paramedical professionals during the antenatal or postnatal period on neonatal mortality rate. a. Forest plot showing the effect of home visitations on the neonatal mortality rate; b. Funnel plot for estimates in meta-analysis for home visits and neonatal mortality rate; c. Forest plot showing the effect of women's group meetings on neonatal mortality rate; d. Funnel plot for estimates in meta-analysis for women's group meetings and neonatal mortality rate.

$I^2 = 0\%$);[34,35,45] and nine trials with WGMs as an intervention (OR 0.68 [95% CI 0.47, 1.00]; $P = 0.05$, $I^2 = 57\%$).[27–29,32,33,43–46] resulted in non-significant decline in MMR (Table 2; Fig. 3a and c).

One study by Pasha et al. could not be pooled because of unavailability of the data needed for calculating OR. It reported a statistically non-significant impact of WGMs on MMR ($P = 0.73$).[47]

Result of the subgroup analysis

Subgroup analysis of trials with WGMs was done based on the proportion of pregnant women attending these meetings. It was observed that NMR (OR 0.67; 95% CI 0.58, 0.77; $P < 0.0001$; $I^2 = 31\%$).[27,28,31,33,36,43,45] and MMR (OR 0.55; 95% CI 0.36, 0.84; $P = 0.005$; $I^2 = 27\%$)[27,28,33,43,45] declined significantly when trials with more than 10% proportion of pregnant women attending WGMs were pooled, respectively. Heterogeneity is also reduced during the subgroup analysis (Fig. 4a and b).

LBW

Three studies with HVs by paramedical professionals as an intervention reported data on LBW.[25,41,42] These studies could not be pooled, as one of the studies was not a cluster randomized trial,[25] and it is not recommended to include only two studies in a meta-analysis. Leroux et al. reported statistically significant reduction in the incidence of LBW ($P < 0.05$).[41] Rotheram et al. reported improvement in infant birth weight due to HVs (OR 0.80 [95% CI 0.57, 1.02], $P = 0.01$).[42] One study reported a significant impact of HVs by paramedical professionals on reducing the incidence of LBW (OR 10.05 [95% CI 3.22, 31.30]; $P < 0.001$].[25]

Stillbirth rate

There were thirteen trials which had SBR as a secondary outcome.[24,26,27,29,30,32–35,43,46,47] The pooled estimates of the five studies with HVs showed a statistically significant reduction in the incidence of stillbirths (OR 0.77, [95% CI 0.70, 0.85]; $P < 0.001$, $I^2 = 0\%$).[24,26,30,34,35] (Fig. 5a). Seven studies with WGMs intervention reported a non-significant reduction in the incidence of stillbirths (OR 0.96 [95% CI 0.84, 1.09]; $P = 0.53$, $I^2 = 37\%$).[28,29,31–33,43,46] (Table 2; Fig. 5c). Study by Pasha et al.[47] could not be pooled. It reported a statistically non-significant impact of WGMs on SBR ($P = 0.62$).

IDR

Sixteen studies evaluated the impact of the interventions on the IDR.[24,27–30,32,33,35,36,38–40,43–45,48] Seven trials had HVs as an intervention, 8 trials had WGMs, and one trial had both the interventions in different study arms. Pooled estimates from eight studies with HVs as an intervention (OR 1.20 [95% CI 0.98, 1.47]; $P = 0.08$, $I^2 = 79\%$);[24,30,35,38–40,45,48] and nine trials with WGMs (OR 1.37 [95% CI 0.99, 1.90]; $P = 0.06$, $I^2 = 59\%$)[27–29,32,33,36,43–45] showed a statistically non-significant improvement in the IDR (Table 2).

Sensitivity analysis

NMR

Sensitivity analysis of the nine studies with HVs as an intervention reported no change in the results when the studies with weights less than 10% were excluded from the analysis.[31,45] The exclusion of a study by Kumar et al., having the combination of HVs and WGMs as intervention, resulted in the reduction of heterogeneity from 77% to 53%.[24] The sensitivity analysis of the studies with WGMs had shown that after excluding More et al. study, the heterogeneity reduced from 71% to 59%.[29] Exclusion of the studies having low weight in the pooling did not change the results.

MMR

Sensitivity analysis of the nine studies with WGM as an intervention has shown no significant change in the pooled results or heterogeneity. Out of three studies with HVs as intervention, Bhutta et al. study weighted 82%.[49] As the number of studies was only three, sensitivity analysis could not be performed for this part.

Stillbirth rate

Excluding the study with less than 10% weight in pooled data for either type of intervention study with HVs or WGMs did not change the overall results.

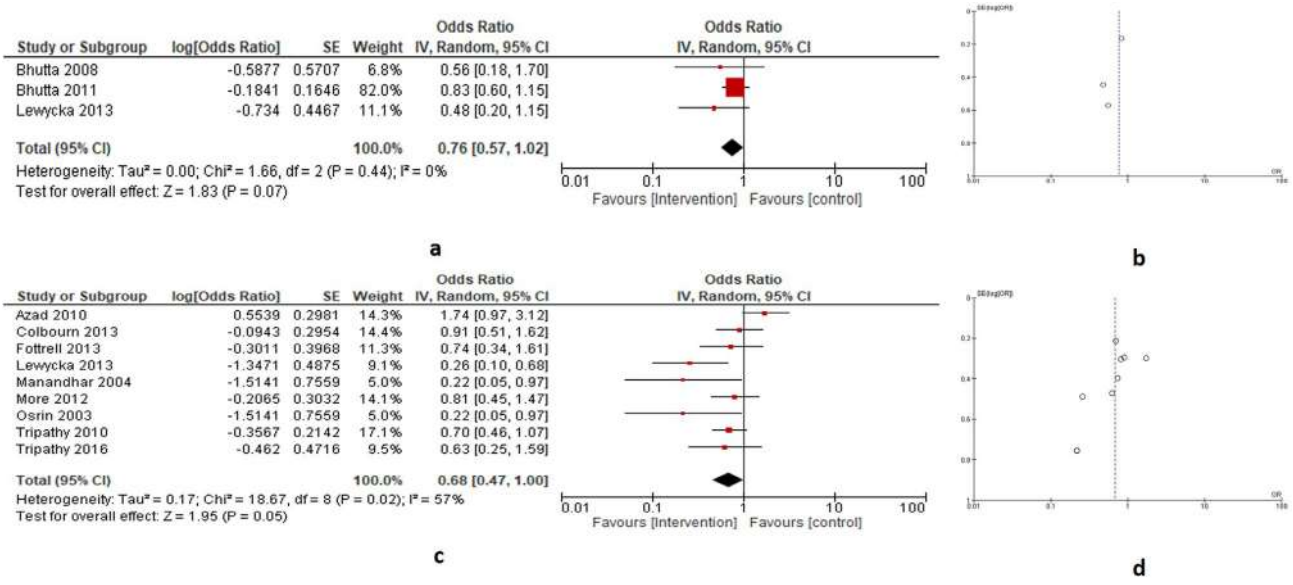


Fig. 3. Meta-analysis of effect of home visits and women's group intervention by paramedical professionals during the antenatal or postnatal period on maternal mortality ratio. a. Forest plot showing the effect of home visitations on the maternal mortality ratio; b. Funnel plot for estimates in meta-analysis for home visits and maternal mortality ratio; c. Forest plot showing the effect of for women's group meetings on the maternal mortality ratio; d. Funnel plot for estimates in meta-analysis for women's group meetings and maternal mortality ratio.

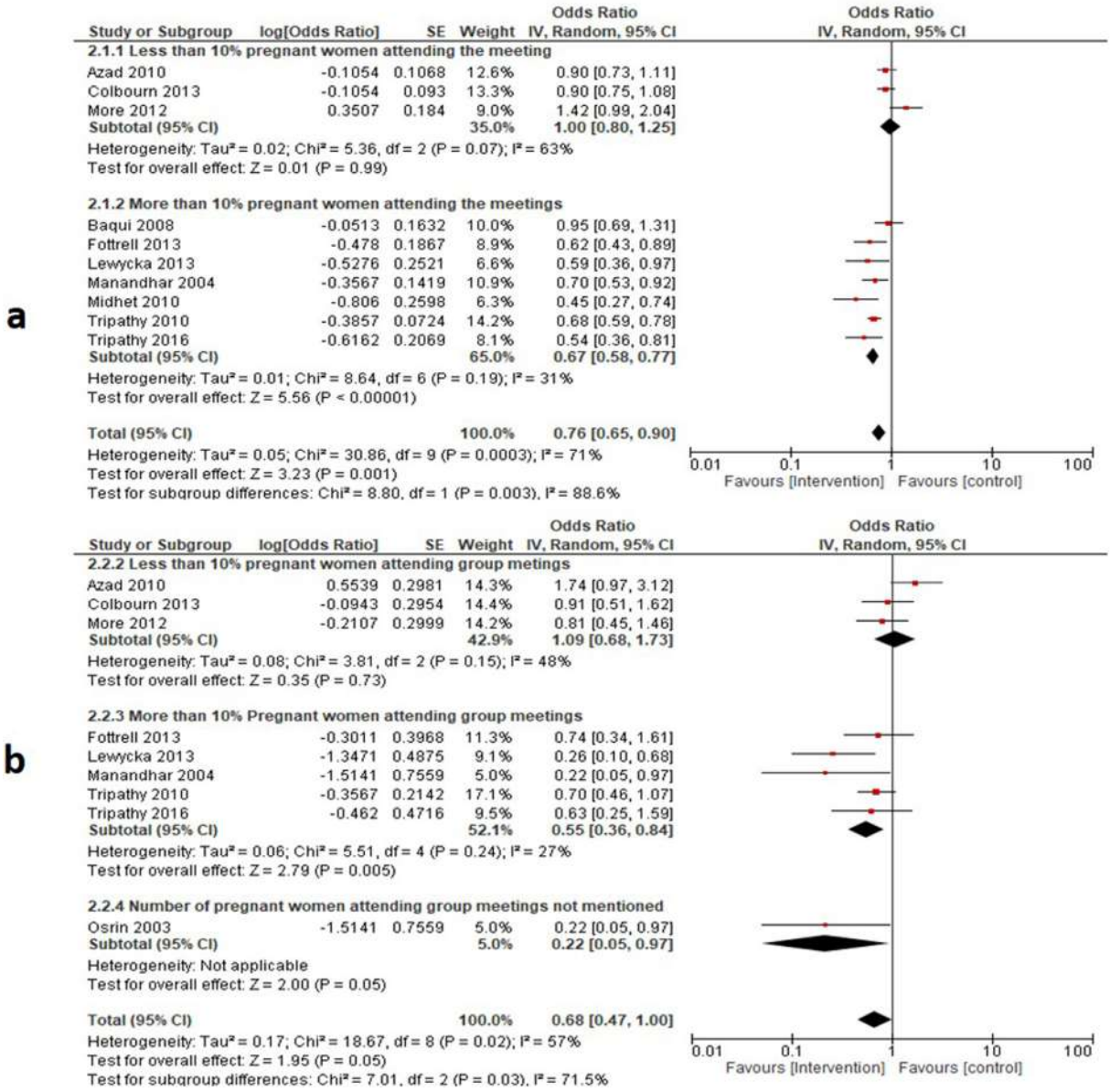


Fig. 4. Subgroup analysis of women's group intervention on a. Neonatal mortality rate b. Maternal mortality ratio and on the basis of proportion of pregnant women attending the meetings.

IDR

Sensitivity analysis of studies with WGMs after excluding studies weighing 5% or less has shown no impact on the pooled results but a significant reduction in the heterogeneity.[43,44] In the group of eight studies with HVs as intervention, the sensitivity analysis reported no significant change in the overall impact on the IDR.

Risk of bias in included studies

Eligible trials had focused more on blinding the outcomes assessors rather than the participants because of the nature of the intervention. Overall, random sequence generation was found to be adequate in 18 trials[24,26,27,29–32,35,38–40,42–46,48] and unclear in seven trials.[28,33,36,37,41,42,47] Twelve trials had reported allocation concealment clearly[24,30,31,33–35,38,42,43,46–48] Others did not report details of allocation.[25–29,32,36,37,39–41,44,45] Five studies[33,35,42,43,45] distinctly reported the process of blinding of data collectors, but in other studies, no information was provided on the same. The risk of bias in the plot has been shown in Fig. 6. The funnel plot was broadly symmetric for the outcomes including NMR, MMR, and SBR (Figs. 2–4).

Quality assessment

Using the GRADE approach, the level of certainty of the evidence was moderate for the evidence generated for HVs as well as WGMs intervention for reducing the NMR and still birth rate.[22] The level of certainty of the evidence was found to be low for improvement in the IDR. The summary of findings as per the GRADE approach, including the reasons for downgrading the evidence generated is given in Table 3.

Discussion

This systematic review and meta-analysis provided evidence on the significant impact of HVs and WGMs by paramedical

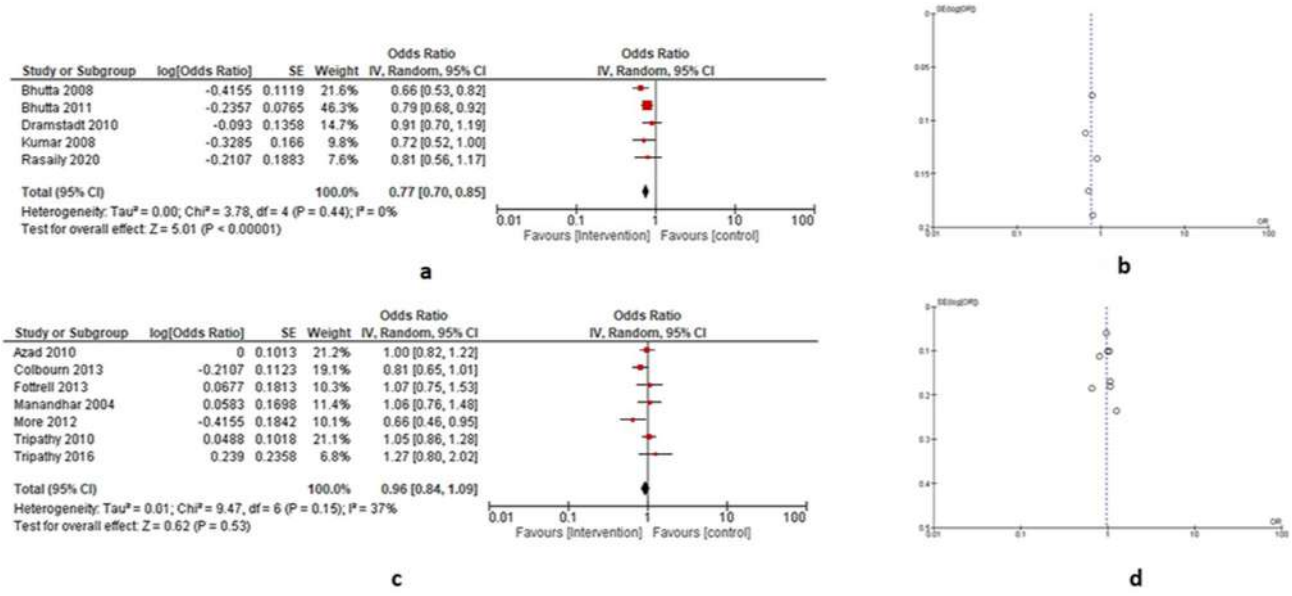


Fig. 5. Meta-analysis of effect of home visits and women's group intervention by paramedical professionals during the antenatal or postnatal period on still birth rate. a. Forest plot showing the effect of home visitations on the still birth rate; b. Funnel plot for estimates in meta-analysis for home visits and still birth rate; c. Forest plot showing the effect of women's group meetings on still birth rate; d. Funnel plot for estimates in meta-analysis for women's group meetings and still birth rate.

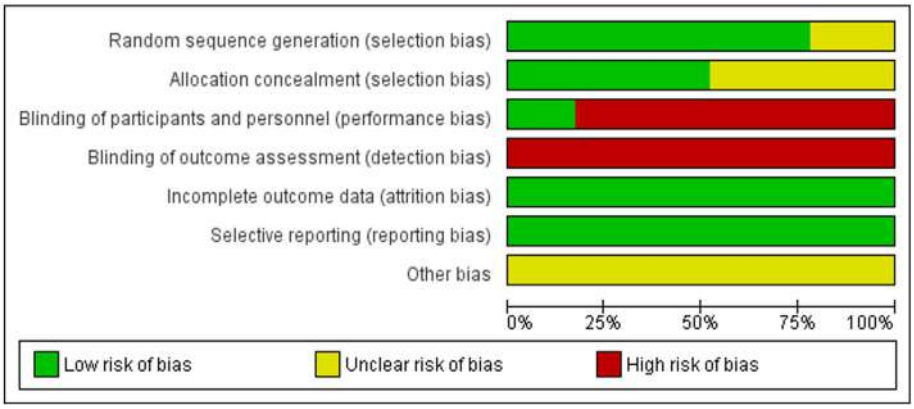


Fig. 6. Risk of bias in studies on home visits or women group meetings by paramedical professionals to improve birth outcomes in low-and-middle-income countries.

professionals in reducing NMR in LMICs. HVs were also effective intervention in reducing the SBR significantly. A significant effect of WGMs in reducing MMR in the subgroup analysis where more than 10% proportion of pregnant women attending the group sessions was observed. LBW incidence and IDR were also reduced by these interventions in the LMICs.

The findings of this review are in line with the previous systematic reviews in low resource settings.[13,14] Prost et al.'s systematic review on WGMs reported 23% reduction in NMR and 37% in MMR, which is similar to 24% reduction in NMR and 32% in MMR observed in this review.[14] Mbuagbaw et al.'s Cochrane's review also reported statistically non-significant impact of community-based interventions alone (OR 0.69, 95% CI 0.45, 1.08, $I^2 = 0\%$) or in combination with health-system interventions (OR 0.70, 95% CI 0.39, 1.26, $I^2 = 0\%$) on MMR.[12] The significant reduction in NMR due to HVs strategy and WGMs observed in this review was found to be similar to Lassi et al. study (OR 0.76 [95% CI 0.68, 0.84] $P < 0.001$, $I^2 = 69\%$).[50]

The results of this review indicate that HVs and WGMs by paramedical professionals were successful in improving NMR. This could be because these interventions strengthened the implementation of the existing MCH programs by mobilizing the community to utilize MCH services by ensuring early registration of pregnancy, providing individualized home-based antenatal care including counseling on diet during antenatal period, birth preparedness, hygiene maintenance during delivery, early initiation of breastfeeding and making them aware regarding danger signs during pregnancy and neonatal period and empowering them to take prompt actions in case of emergency which had significant impact on reducing still births and NMR.

Both HVs and WGMs interventions have shown reduction in MMR as well, although it was statistically non-significant. A possible reason for this might be that there is limited number of studies that had HVs as an intervention, and WGMs generally target women in the reproductive age group with a small proportion of pregnant women attending these meetings. The results of the subgroup analysis showed 45% decline in MMR (OR 0.55; 95% CI 0.36, 0.846; $P = 0.005$, $I^2 = 27\%$) in the pooled estimates of the studies where the participation of pregnant women was more than 10% in the WGMs (Fig. 5). Prost et al.'s

Table 3
GRADE summary of findings.

a) Impact of home visits by paramedical professionals on various birth outcomes

Certainty assessment							Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Relative 95% CI	Absolute 95% CI		
Neonatal mortality rate										
9	randomised trials	serious[a]	not serious	not serious	not serious	none	OR 0.77 (0.67 to 0.90)	1 fewer per 1,000 1,000 (from 1 fewer to 1 fewer)	⨁⨁⨁◯ MODERATE	IMPORTANT
Maternal mortality ratio										
3	randomised trials	serious[a]	not serious	not serious	serious[b]	none	OR 0.76 (0.57 to 1.02)	1 fewer per 1,000 1,000 (from 1 fewer to 1 fewer)	⨁⨁◯◯ LOW	IMPORTANT
Still birth rate										
5	randomised trials	serious[a]	not serious	not serious	not serious	none	OR 0.77 (0.70 to 0.85)	1 fewer per 1,000 1,000 (from 1 fewer to 1 fewer)	⨁⨁⨁◯ MODERATE	IMPORTANT
Institutional delivery rate										
8	randomised trials	serious[a]	not serious	serious[a,c]	not serious	none	OR 1.20 (0.98 to 1.47)	1 fewer per 1,000 1,000 (from 1 fewer to 1 fewer)	⨁⨁◯◯ LOW	CRITICAL

b) Impact of women group meetings on various birth outcomes

Certainty assessment							Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Relative 95%CI	Absolute 95%CI		
Neonatal Mortality Rate										
10	randomised trials	serious[a]	not serious	not serious	not serious	none	**OR 0.76 (0.65 to 0.90)**	**1 fewer per 1,000 (from 1 fewer to 1 fewer)**	⨁⨁⨁◯ MODERATE	IMPORTANT
Maternal Mortality Ratio										
9	randomised trials	serious[a]	not serious	not serious	not serious	none	**OR 0.68 (0.47 to 1.00)**	**1 fewer per 1,000 (from 1 fewer to 1 fewer)**	⨁⨁⨁◯ MODERATE	IMPORTANT
Stillbirth Rate										
8	randomised trials	serious[a]	not serious	not serious	not serious	none	**OR 0.96 (0.84 to 1.09)**	**1 fewer per 1,000 (from 1 fewer to 1 fewer)**	⨁⨁⨁◯ MODERATE	IMPORTANT
Number of institutional deliveries										
9	randomised trials	serious[a]	not serious	serious[b]	not serious	none	**OR 1.20 (0.98 to 1.47)**	**1 fewer per 1,000 (from 1 fewer to 1 fewer)**	⨁⨁◯◯ LOW	IMPORTANT

CI: Confidence interval; **OR:** Odds ratio
Explanations
[a] Studies had been conducted in different geographic regions with participants having different socio-demographic characteristics. Blinding of the participants and outcome assessors was not feasible in majority of the study settings.
[b] Majority of the studies had not reported institutional deliveries as an outcome. The impact of the intervention could not be assessed due to different strategies of the health programs in different health settings.
[c] Inadequate number of studies. Hence the impact of intervention could not be generalized.

systematic review had reported a 55% decline in MMR in the groups where the proportion of pregnant women attending WGMs was more than 30% (OR 0.45; 95% CI 0.17–0.73; $P = 0.104$, $I^2 = 51.3\%$).[14] This indicated that targeting pregnant women in WGMs is more important than women in reproductive age group as a whole. Kidney et al.'s, systematic review included studies with all types of community-based interventions (prenatal HVs, WGMs, delivery of sterile kits).[12] It had reported a reduction of 38% in the MMR (OR 0.62; 95% CI 0.39, 0.98; $P = 0.04$), which indicated that community-based interventions such as WGMs, training of birth attendants, provision of sterile delivery kits, and so on are more effective in lowering the MMR.

The evidence was found to be low for outcomes like IDR as the trials included in the review were designed to measure impact on NMR and not on IDR. Although the studies with HVs by paramedical professionals and WGMs were pooled separately, we found considerable statistical heterogeneity (77% and 71% for NMR, 0% and 57% for MMR, 0% and 37% for SBR, and 79% and 59% for IDR, respectively). The probable reason for this heterogeneity could be due to differences in the characteristics of the study populations and health interventions and infrastructure. Heterogeneity reduced after removing outlier studies.

The strengths of this review are the use of broad definition of community-based interventions, a comprehensive search of the literature with a wide range of databases, hand-searching of gray literature, GRADE approach, and level of certainty of the evidence.[22] There was a minimal risk of publication bias as the funnel plot (Figs. 2, 3 and 5) was broadly symmetric for all the outcomes, so the evidence was not downgraded further.

The impact of the individual interventions such as imparting knowledge for birth preparedness,[24,31,35,38,39] education on nutrition and diet during pregnancy[25,29,41,42] or promoting exclusive breastfeeding,[24,38,41,45] and felicitated referral of the pregnant woman/neonate[30,38,47] in case needed was not studied because of the non-availability of intervention specific data. This could be treated as a gap in the study and could provide basis for further research as these interventions might have contributed in reduction of the NMR, MMR, SBR, LBW, and IDR.

The public health implications of the findings of this systematic review are that as HVs and WGMs interventions were significantly effective in reducing the NMR; HVs in reducing the SBR; and higher participation of pregnant women in the WGMs in reducing MMR. The studies with HVs as an intervention reported an average of three antenatal and four postnatal visits at home. The antenatal visits started during the first trimester of the pregnancy and focused primarily on the identification of high-risk pregnancies, promoting healthy diet and hygiene during the pregnancy. The CHWs were given financial incentives for motivating the woman for opting institutional or skilled birth delivery. The CHWs gave postnatal visits till six weeks after the delivery focusing primarily on newborn care including counseling on early initiation and exclusive breastfeeding, delaying first bath of the child and skin to skin care of the baby and special care for low-birth-weight babies. The postnatal visits also included the identification of early danger signs and referral (if needed) in the newborn as well as educating the mother for the same. Participatory learning approach of WGMs focused on women in reproductive age group. The WGMs were held either monthly or fortnightly. The intervention package included promoting healthy diet and hygiene during pregnancy, birth preparedness and motivating the woman to utilize health facilities for ANC and delivery. Follow-up meetings with the same women groups were held next month so as to maintain contact with the woman as well as for identification and registration of pregnancy in the first trimester itself. The pregnant ladies were then counseled to utilize the health facilities during antenatal period. A study by Kumar et al. reported better results in terms of reduction in NMR (OR 0.46, 95% CI 0.35–0.60) when HVs were complemented by WGMs in the community compared to the studies with individual interventions.[24] The community health interventions are effective in countries with scarcity of health workforce. Obviously, they can not replace the trained health workforce but they have the potential to support them resulting in expanding access and utilization of health services. If provided adequate training, equipment, and supportive supervision by the existing health system, these CHWs can act as a bridge between the community and the health system. Hence, LMICs should strengthen and continue implementing both these interventions to further reduce the NMR, MMR, and SBR to achieve the SDGs related to MCH.

Author statements

Acknowledgments

We would like to thank Dr. Rajesh Kumar, Professor, and Head, Department of Community Medicine and Dean (Academics), Postgraduate Institute of Medical Education and Research, Chandigarh, India in providing his expert guidance during manuscript preparation, and Dr. Meenu Singh, Professor, Advanced Paediatric Center in guiding and training in conducting systematic review and meta-analysis.

Ethical approval

None required.

Funding

We did not receive any funding for conducting this review. Dr. Chakrapani was in part supported by the Wellcome Trust/DBT India Alliance Senior Fellowship (IA/CPHS/16/1/502667).

Competing interests

None declared.

Author contributions

- Madhu Gupta, MD, PhD: She conceptualized the study, reviewed the studies, analyzed the data, provided intellectual inputs to improve the quality of the manuscript, and gave the final approval to submit the manuscript
- Adarsh Bansal, MPH: He has written the first draft of the manuscript, reviewed the studies, analyzed the data, and gave the final approval to submit the manuscript
- Venkatesan Chakrapani, MD: Reviewed the studies, analyzed the data, provided intellectual inputs to improve the quality of the manuscript, and gave the final approval to submit the manuscript
- Nishant Jaiswal, PhD: Reviewed the studies, analyzed the data, provided intellectual inputs to improve the quality of the manuscript, and gave the final approval to submit the manuscript
- Tanvi Kran, PhD: analyzed the data, provided intellectual inputs to improve the quality of the manuscript, and gave the final approval to submit the manuscript

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2022.11.023.

References

1. World Health Organization. *Sustainable development goals*. Available from: https://www.un.org/sustainabledevelopment/sustainable-development-goals/, 2015. [Accessed 17 January 2018].
2. World Health organization. *Children: reducing mortality*. Available at: http://www.who.int/news-room/fact-sheets/detail/children-reducing-mortality, 2017. [Accessed 17 January 2018].
3. UNICEF, WHO, World Bank, UN-DESA Population Division. *Levels and trends in child mortality 2015*. World Health Organization; 2015. Available at: http://www.who.int/maternal_child_adolescent/documents/levels_trends_child_mortality_2015/en/. [Accessed 29 August 2018].
4. Million Death Study Collaborators. Causes of neonatal and child mortality in India: a nationally representative mortality survey. *Lancet* 2010;**376**:1853–60.
5. World Health Organization. *Newborn death and illness*. Available from: http://www.who.int/pmnch/media/press_materials/fs/fs_newborndealth_illness/en/, 2011. [Accessed 17 January 2018].
6. Callister L, Edwards J. Sustainable development goals and the ongoing process of reducing maternal mortality. *J Obstet Gynecol Neonatal Nurs* 2017;**46**:e56–64.
7. Brenner J, Kabakyenga J, Kyomuhangi T, Wotton K, Pim C, Ntaro M, et al. Can volunteer community health workers decrease child morbidity and mortality in southwestern Uganda? An impact evaluation. *PLoS One* 2011;**6**(12):e27997.
8. Ministry of Health. Government of Ethiopia. Health extension and education center. Health extension program in Ethiopia. In: *Federal Ministry of health*; 2007. http://www.moh.gov.et/english/Resources/Documents/HEW%20profile%20Final%2008%2007.pdf. [Accessed 10 May 2021].
9. Hafeez A, Mohamud B, Shiekh M, Shah S, Jooma R. Lady health workers programme in Pakistan: challenges, achievements and the way forward. *J Pak Med Assoc* 2011;**61**(3):210.
10. National Health Mission. *Ministry of health and family welfare*. Government of India; 2021. Accessed from: https://nhm.gov.in/index1.php?lang=1&level=1&sublinkid=150&lid=226. Accessed on 10.05.2021.
11. Barger D, Owen H, Pitt C, Kerber K, Sitrin D, Mayora C, et al. Multi-country analysis of the cost of community health workers kits and commodities for

community-based maternal and newborn care. *Health Policy Plann* 2017;**32**(suppl_1):i84–92. Oct 1.
12. Kidney E, Winter H, Khan K, Gülmezoglu A, Meads C, Deeks J, et al. Systematic review of effect of community-level interventions to reduce maternal mortality. *BMC Pregnancy Childbirth* 2009;**9**:2.
13. Mbuagbaw L, Medley N, Darzi A, Richardson M, Habiba Garga K, Ongolo-Zogo P. Health system and community level interventions for improving antenatal care coverage and health outcomes. *Cochrane Database Syst Rev* 2015:CD010994.
14. Prost A, Colbourn T, Seward N, Azad K, Coomarasamy A, Copas A, et al. Women's groups practising participatory learning and action to improve maternal and newborn health in low-resource settings: a systematic review and meta-analysis. *Lancet* 2013;**381**:1736–46.
15. Liberati A, Tetzlaff J, Mulrow C, Gøtzsche P, Ioannidis J, Clarke M, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;**6**:e1000100.
16. Gupta M, Bansal A, Chakrapani V, Kumar R. *Effectiveness of prenatal and postnatal home visits by paramedical professionals in improving maternal and child health outcomes in developing countries: a systematic review of intervention studies*. Available at: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018091968; 2018. Prospero.
17. Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M, et al. *Searching for and selecting studies. Cochrane handbook for systematic reviews of interventions*. 2019. p. 67–107.
18. Fantom N, Serajuddin U. *The World Bank's classification of countries by income*. The World Bank; 2016.
19. Fernandez P. *Zotero*. 2013. Available from: https://www.zotero.org. [Accessed 17 January 2018].
20. Higgins J, Altman D, Gøtzsche P, Jüni P, Moher D, Oxman A, et al. The Cochrane collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;**343**:d5928.
21. Lopez-Vazquez P, Vazquez-Lago JM, F A. Misprescription of antibiotics in primary care: a critical systematic review of its determinants. *J Eval Clin Pract* 2012;**18**:473–84.
22. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;**7**:177–88.
23. Alonso-Coello P, Schünemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ* 2016;**353**:i2016.
24. Kumar V, Mohanty S, Kumar A, Misra R, Santosham M, Awasthi S, et al. Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial. *Lancet* 2008;**372**:1151–62.
25. Singh R, Shaikh S, Hassan M, Dwivedi S. Effect of ensured ante-natal care on birth outcome: a community-based interventional study. *Health Popul Perspect Issues* 2014;**37**:88–97.
26. Rasaily R, Saxena NC, Pandey S, Garg BS, Swain S, Iyengar SD, et al. Effect of home-based newborn care on neonatal and infant mortality: a cluster randomised trial in India. *BMJ Glob Health* 2020;**5**:e000680.
27. Tripathy P, Nair N, Barnett S, Mahapatra R, Borghi J, Rath S, et al. Effect of a participatory intervention with women's groups on birth outcomes and maternal depression in Jharkhand and Orissa, India: a cluster-randomised controlled trial. *Lancet* 2010;**375**:1182–92.
28. Tripathy P, Nair N, Sinha R, Rath S, Gope R, Rath S, et al. Effect of participatory women's groups facilitated by Accredited Social Health Activists on birth outcomes in rural eastern India: a cluster-randomised controlled trial. *Lancet Glob Health* 2016;**4**:e119–28.
29. More N, Bapat U, Das S, Alcock G, Patil S, Porel M, et al. Community mobilization in Mumbai slums to improve perinatal care and outcomes: a cluster randomized controlled trial. *PLoS Med* 2012;**9**:e1001257.
30. Darmstadt G, Choi Y, Arifeen S, Bari S, Rahman S, Mannan I, et al. Evaluation of a cluster-randomized controlled trial of a package of community-based maternal and newborn interventions in Mirzapur, Bangladesh. *PLoS One* 2010;**5**:e9696.
31. Baqui A, El-Arifeen S, Darmstadt G, Ahmed S, Williams E, Seraji H, et al. Effect of community-based newborn-care intervention package implemented through two service-delivery strategies in Sylhet district, Bangladesh: a cluster-randomised controlled trial. *Lancet* 2008;**371**:1936–44.
32. Azad K, Barnett S, Banerjee B, Shaha S, Khan K, Rego A, et al. Effect of scaling up women's groups on birth outcomes in three rural districts in Bangladesh: a cluster-randomised controlled trial. *the Lancet* 2010;**375**:1193–202.
33. Fottrell E, Azad K, Kuddus A, Younes L, Shaha S, Nahar T, et al. The effect of increased coverage of participatory women's groups on neonatal mortality in Bangladesh: a cluster randomized trial. *JAMA Pediatr* 2013;**167**:816–25.
34. Bhutta Z, Memon Z, Soofi S, Salat M, Cousens S, Martines J. Implementing community-based perinatal care: results from a pilot study in rural Pakistan. *Bull World Health Organ* 2008;**86**:452–9.
35. Bhutta Z, Soofi S, Cousens S, Mohammad S, Memon Z, Ali I, et al. Improvement of perinatal and newborn care in rural Pakistan through community-based strategies: a cluster-randomised effectiveness trial. *Lancet* 2011;**377**:403–12.
36. Midhet F, Becker S. Impact of community-based interventions on maternal and neonatal health indicators: results from a community randomized trial in rural Balochistan, Pakistan. *Reprod Health* 2010;**7**:30.
37. Memon Z, Khan G, Soofi S, Baig I, Bhutta Z. Impact of a community-based perinatal and newborn preventive care package on perinatal and neonatal mortality in a remote mountainous district in Northern Pakistan. *BMC Pregnancy Childbirth* 2015;**15**:106.
38. Penfold S, Manzi F, Mkumbo E, Temu S, Jaribu J, Shamba D, et al. Effect of home-based counselling on newborn care practices in southern Tanzania one year after implementation: a cluster-randomised controlled trial. *BMC Pediatr* 2014;**14**:187.
39. Hanson C, Manzi F, Mkumbo E, Shirima K, Penfold S, Hill Z, et al. Effectiveness of a home-based counselling strategy on neonatal care and survival: a cluster-randomised trial in six districts of rural southern Tanzania. *PLoS Med* 2015;**12**:e1001881.
40. Geldsetzer P, Mboggo E, Larson E, Lema I, Magesa L, Machumi L, et al. Community health workers to improve uptake of maternal healthcare services: a cluster-randomized pragmatic trial in Dar es Salaam, Tanzania. *PLoS Med* 2019;**16**.
41. Le Roux I, Rotheram-Borus M, Stein J, Tomlinson M. The impact of paraprofessional home visitors on infants' growth and health at 18 months. *Vulnerable Child Youth Stud* 2014;**9**:291–304.
42. Rotheram-Borus M, Tomlinson M, Le Roux I, Harwood J, Comulada S, O'Connor M, et al. A cluster randomised controlled effectiveness trial evaluating perinatal home visiting among South African mothers/infants. *PLoS One* 2014;**9**:e105934.
43. Manandhar D, Osrin D, Shrestha B, Mesko N, Morrison J, Tumbahangphe K, et al. Effect of a participatory intervention with women's groups on birth outcomes in Nepal: cluster-randomised controlled trial. *Lancet* 2004;**364**:970–9.
44. Osrin D, Mesko N, Shrestha B, Shrestha D, Tamang S, Thapa S, et al. Implementing a community-based participatory intervention to improve essential newborn care in rural Nepal. *Trans R Soc Trop Med Hyg* 2003;**97**:18–21.
45. Lewycka S, Mwansambo C, Rosato M, Kazembe P, Phiri T, Mganga A, et al. Effect of women's groups and volunteer peer counselling on rates of mortality, morbidity, and health behaviours in mothers and children in rural Malawi (MaiMwana): a factorial, cluster-randomised controlled trial. *Lancet* 2013;**381**:1721–35.
46. Colbourn T, Nambiar B, Bondo A, Makwenda C, Tsetekani E, Makonda-Ridley A, et al. Effects of quality improvement in health facilities and community mobilization through women's groups on maternal, neonatal and perinatal mortality in three districts of Malawi: MaiKhanda, a cluster randomized controlled effectiveness trial. *Int Health* 2013;**5**:180–95.
47. Pasha O, McClure E, Wright L, Saleem S, Goudar S, Chomba E, et al. A combined community-and facility-based approach to improve pregnancy outcomes in low-resource settings: a Global Network cluster randomized trial. *BMC Med* 2013;**11**:215.
48. Kirkwood B, Manu A, ten Asbroek A, Soremekun S, Weobong B, Gyan T, et al. Effect of the Newhints home-visits intervention on neonatal mortality rate and care practices in Ghana: a cluster randomised controlled trial. *Lancet* 2013;**381**:2184–92.
49. Bhutta Z, Lassi Z, Pariyo G, Huicho L. *Global experience of community health workers for delivery of health related millennium development goals: a systematic review, country case studies, and recommendations for integration into national health systems. Global Health Workforce Alliance*, vol. 1; 2010. p. 61.
50. Lassi Z, Haider B, Bhutta Z. Community-based intervention packages for reducing maternal and neonatal morbidity and mortality and improving neonatal outcomes. *Cochrane Database Syst Rev* 2010;**11**.

Public Health 215 (2023) 118–123

Contents lists available at ScienceDirect

Public Health

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

Original Research

The impact of state paid sick leave policies on weekday workplace mobility during the COVID-19 pandemic

C.C. Pollack [a, b, c, *], A. Deverakonda [a], F. Hassan [a, d], S. Haque [a, e], A.N. Desai [a, f], M.S. Majumder [a, g, h]

[a] COVID-19 Dispersed Volunteer Research Network; Boston, Massachusetts, USA
[b] Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth College; Lebanon, New Hampshire, USA
[c] Department of Epidemiology, Geisel School of Medicine at Dartmouth College; Lebanon, New Hampshire, USA
[d] School of Public Health, University of Alberta; Edmonton, Alberta, Canada
[e] Network Science Institute, Northeastern University; Boston, Massachusetts, USA
[f] Department of Internal Medicine, Division of Infectious Diseases, University of California - Davis; Sacramento, California, USA
[g] Computational Health Informatics Program, Boston Children's Hospital; Boston, Massachusetts, USA
[h] Department of Pediatrics, Harvard Medical School; Boston, Massachusetts, USA

ARTICLE INFO

Article history:
Received 9 May 2022
Received in revised form
12 August 2022
Accepted 26 August 2022
Available online 12 September 2022

Keywords:
COVID-19
Paid sick leave
Physical distancing
Workplace mobility
Health policy

ABSTRACT

Objectives: This study aimed to evaluate whether the Families First Coronavirus Response Act (FFCRA) modified the association between pre-existing state paid sick leave (PSL) and weekday workplace mobility between February 15 and July 7, 2020.
Study design: This was a longitudinal, observational study.
Methods: The 50 US states and Washington, D.C., were divided into exposure groups based on the presence or absence of pre-existing state PSL policies. Derived from Google COVID-19 Community Mobility Reports, the outcome was measured as the daily percent change in weekday workplace mobility. Mixed-effects, interrupted time series regression was performed to evaluate weekday workplace mobility after the implementation of the FFCRA on April 1, 2020.
Results: States with pre-existing PSL policies exhibited a greater drop in mobility following the passage of the FFCRA ($\beta = -8.86$, 95% confidence interval: −11.6, −6.10, $P < 001$). This remained significant after adjusting for state-level health, economic, and sociodemographic indicators ($\beta = -3.13$, 95% confidence interval: −5.92, −0.34; $P = .039$).
Conclusions: Pre-existing PSL policies were associated with a significant decline in weekday workplace mobility after the FFCRA, which may have influenced local health outcomes. The presence of pre-existing state policies may differentially influence the impact of federal legislation enacted during emergencies.

Introduction

The COVID-19 pandemic necessitates systemic policies to reduce its spread. Despite the deployment of COVID-19 vaccines, the ability to quarantine after exposure remains critical to minimize the potential for "breakthrough cases" and the risk of infection for those who are unvaccinated.[1] One policy to facilitate self-quarantine is paid sick leave (PSL), which allows employees to take compensated time off from work to recover from illness or injury. PSL has previously been associated with a three-fold increase in the protection of workers' jobs, income, and health while recovering from illness.[2] PSL is especially crucial during outbreaks of communicable diseases, as it can help mitigate "presenteeism," whereby employees go to work even if they are sick.[3] This is particularly important for COVID-19 since individuals can present a range of symptoms.

Although previous studies have shown the efficacy of PSL in reducing absenteeism, these studies have focused on European countries with robust PSL schemes.[4] The United States is one of only two Organisation for Economic Co-operation and Development countries that does not have a nationwide PSL policy, resulting in a patchwork system that varies between states.[2,5] In

* Corresponding author. Rubin Building 833; 1 Medical Center Drive; Lebanon, New Hampshire, USA. Tel.: +1540 497 3419.
E-mail address: catherine.c.pollack.gr@dartmouth.edu (C.C. Pollack).

https://doi.org/10.1016/j.puhe.2022.08.019
0033-3506/

addition, previous studies on PSL and absenteeism in the United States have focused on specific states or localities rather than taking a national approach.[6,7] Within each state, access to PSL is associated with many factors, including industry type, race, ethnicity, gender, sexual orientation, income level, immigration status, company size, full-time or part time status, and experience level. As a result, up to 40% of American private sector workers, including 69% of the lowest quartile of wage earners, are not afforded PSL.[8] This was partially rectified with the Families First Coronavirus Response ACT (FFCRA) and Coronavirus Aid, Relief and Economic Security Act, which provided emergency, 2-week PSL on April 1, 2020.[9] This federally legislated PSL played an important role in slowing the spread of COVID-19 in the workplace by allowing for self-quarantine from work environments.[9–11] However, exemptions for certain employee categories (e.g. healthcare workers and emergency responders) and businesses with more than 500 employees blunted its coverage to potentially as few as 47% of private-sector workers.[10] Thus, the presence of pre-existing state PSL may have influenced how this emergency federal legislation impacted key outcomes, such as travel to and from the workplace (i.e. weekday workplace mobility), which could be considered a proxy for workplace presenteeism and absenteeism.[11,12] As a result, it is critical to identify the differential impacts of the FFCRA on states that had pre-existing state PSL to elucidate what fundamental level of local preparedness is required to maximize the impact of federal legislation. The purpose of this study was to explore the impact of pre-existing state PSL on weekday workplace mobility surrounding the passage of the FFCRA (i.e. February to July 2020). It was hypothesized that states that had pre-existing state PSL would experience a greater drop in weekday workplace mobility compared with states that did not.

Methods

Data collection

Four data sets were integrated for each of the 50 states and Washington, DC. The primary exposure of interest (i.e. presence or absence of pre-existing state PSL) was coded as either "yes" or "no" based on data from the Kaiser Family Foundation.[5] The primary outcome of interest (i.e. weekday workplace mobility) was collected from Google COVID-19 Community Mobility Reports.[13] Within these reports, weekday workplace mobility was calculated as the percent change in mobility between the date of interest and a prepandemic baseline. This baseline was computed as the median mobility between January 3 and February 6, 2020, on the same day of the week (e.g. Monday, Tuesday) as the date of interest. Economic covariates (e.g. wage policies, worker protection policies, right-to-organize policies) and epidemiological metrics (e.g. COVID-19 cases and deaths per state) were from the Oxfam Index and the *New York Times* COVID-19 database, respectively. Other sociodemographic factors (e.g. median household income, state gross domestic product, commuting patterns, presidential election results between 2004 and 2016) were from the American Community Survey and the Federal Election Commission.[14–17]

Statistical analysis

A mixed-effects, interrupted time series regression model with nested random effects for state and month characterized the relationship between the presence of pre-existing state PSL and daily percent change in weekday workplace mobility. The initial model only adjusted for temporality relative to the implementation of the FFCRA on April 1, 2020 (i.e. days pre-FFCRA, instantaneous FFCRA, and days post-FFCRA). Additional bivariate analyses were performed to identify which covariates were significantly associated with weekday workplace mobility. Highly correlated terms were evaluated by investigators to determine which should be retained for further analysis. A multivariable model was subsequently constructed with the same structure as the unadjusted model and all significant terms from the bivariate analysis. Data were aggregated with Python (version 3.8) and analyzed in R (version 4.0.3) using the RStudio Integrated Development Environment (version 1.3.1093).

Results

Immediately after FFCRA implementation on April 1, 2020, Washington DC and the 12 states with pre-existing state PSL experienced an 8.86 percentage point greater decrease in weekday workplace mobility ($\beta = -8.86$, 95% confidence interval CI: −11.6, −6.10, $P < .001$) compared with the 39 states that do not have pre-existing state PSL (Fig. 1). The substantial drop in weekday workplace mobility before the FFCRA coincided with state-mandated stay-at-home orders. Health indicators associated with a greater decrease in mobility included new cases per 100,000 ($\beta = -0.03$, 95% CI: −0.04, −0.03; $P < .001$) and new deaths per 100,000 ($\beta = -0.43$, 95% CI: −0.51, −0.35; $P < .001$). Many travel metrics were associated with weekday workplace mobility, although directionality varied. For example, although average commute time was inversely associated with weekday workplace mobility (β per minute = −1.04, 95% CI: −1.22, −0.86; $P < .001$), percent commuting via carpool was associated with an increase in weekday workplace mobility ($\beta = 1.73$, 95% CI: 0.63, 2.83; $P = .003$). The bulk of economic indicators were also associated with weekday workplace mobility, including 2017 median household income (β per $10,000 USD = −2.47, 95% CI: −3.64, −1.29; $P < .001$) and unemployment rate ($\beta = -0.31$, 95% CI: −0.40, −0.20; $P < .001$). In addition, states with a dominant labor sector in "education and health services" had a greater drop in weekday workplace mobility compared with states with a dominant labor sector in "trade, transportation, and utilities" ($\beta = -4.90$, 95% CI: −9.39, −0.42, $P = .044$). Several demographic indicators were also associated with weekday workplace mobility, albeit in various directions. For example, although a higher percentage of men was associated with an increase in weekday workplace mobility ($\beta = 2.83$, 95% CI: 1.11, 4.55; $P = .002$), a higher percentage of Asian individuals was associated with a greater decrease in weekday workplace mobility ($\beta = -0.31$, 95% CI: −0.58, −0.05; $P = .024$). In terms of policies, states that provided paid *family* leave had a greater drop in weekday workplace mobility compared with states that did not ($\beta = -10.6$, 95% CI: −14.8, −7.02; $P < .001$). Finally, a higher state population per square mile was associated with a greater drop in weekday workplace mobility (β per 1000 persons = −2.04, 95% CI: −2.84, −1.23; $P < .001$). Supplementary Table 1 provides comprehensive list of covariates.

After adjustment, the association between pre-existing state PSL and weekday workplace mobility remained statistically significant ($\beta = -3.13$, 95% CI: −5.92, −0.34; $P = .039$; Table 1). Other variables that retained their significance and associated with a decrease in weekday workplace mobility included new cases per 100,000 ($\beta = -0.03$, 95% CI: −0.04, −0.03; $P < .001$), average commute time (β per minute = −0.59, 95% CI: −0.94, −0.24; $P = .004$), unemployment rate ($\beta = -0.35$, 95% CI: −0.45, −0.26; $P < .001$), and state population per square mile (β per 1000 persons = −1.12, 95% CI: −2.04, −0.20; $P = .027$). Variables that retained their significance and were associated with an increase in weekday workplace mobility included poverty rate ($\beta = 0.50$, 95% CI: 0.07, 0.94; $P = .035$) and "manufacturing" as a dominator labor sector relative to "trade, transportation, and utilities" ($\beta = 7.34$, 95% CI: 0.59, 14.1; $P = .045$).

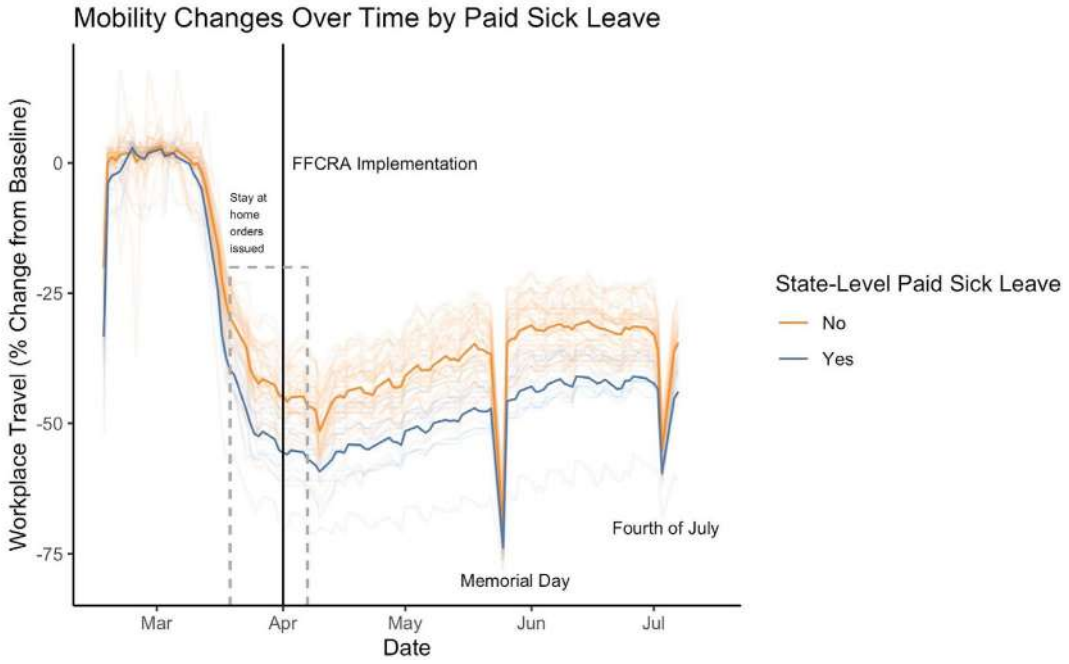


Fig. 1. Changes in workplace travel over time by state-level paid sick leave. The black line on April 1, 2020, denotes the implementation of the Families First Coronavirus Response Act (FFCRA). The gray dashed lines signify the period in which stay-at-home orders were enacted by states. Twelve states (Arizona, California, Connecticut, the District of Columbia, Massachusetts, Maryland, New Jersey, New York, Oregon, Rhode Island, Vermont, and Washington) had pre-existing paid sick leave policies mandated by the state, whereas the remaining 39 did not. The prominent blue and orange lines denote group-level daily averages, whereas the lighter lines are for each individual state. The most substantial drops occurred on two federal US holidays: Memorial Day (May 25, 2020) and Independence Day (July 4, 2020). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Discussion

This study is the first to comprehensively evaluate the impact of pre-existing state PSL on weekday workplace mobility in the United States during the COVID-19 pandemic. The presence of pre-existing state PSL was significantly associated with a drop in weekday workplace mobility in the early phase of the pandemic in both unadjusted and adjusted models. These results suggest a complex interplay between pre-existing labor workforce protections and emergency public health interventions targeted for the workforce.

Increasingly, states are held responsible for managing and administering social services, leading to highly variable policies.[18] The presence of pre-existing state PSL acted as a "classifier" that could differentiate how the FFCRA impacted state weekday workplace mobility. As one of the first major nationwide COVID-19 policies, the impact of any single part of the FFCRA was unprecedented, and the period between the announcement of the legislation and its implementation was relatively short. Coupled with the diverse array of state-level policies that were enacted during this time, it is likely that anticipatory behavior did not substantially influence the observed association between pre-existing state PSL and weekday workplace mobility.

Given the ubiquity of COVID-19, this nationwide, ecological evaluation may suggest that federal emergency aid packages have a stronger impact in localities with the pre-existing infrastructure to support such policies. This study also contributes to the literature characterizing the impact of the FFCRA and its emergency PSL on various health and behavioral outcomes. A prior study, which relied on cellular data in place of Google COVID-19 Community Mobility Reports, also found that the FFCRA significantly decreased the time spent away from home. However, the FFCRA's impact on *workplace* mobility, as is the focus of this study, could not be determined.[12]

As COVID-19 variants of concern continue to emerge, the lack of consistent PSL policies across the United States leaves employees vulnerable, especially those considered "essential workers" or in positions that require in-person work.[19] This disproportionately impacts Black, Indigenous, People of Color, as well as the socio-economically disadvantaged—the same groups that are both at higher risk for COVID-19 and disenfranchised by current labor laws.[20] To protect such individuals, there is a need for permanent structural changes in labor protection laws at the federal level, which could leverage pre-existing state policies to identify best practices and potential pitfalls.[21] Our work also supports similar conclusions regarding PSL schemes in Europe: different levels of labor protection laws correspond to different levels of PSL-supported work absences, underscoring the need for strong, long-term policy support for PSL in both the United States and Europe.[22] Furthermore, systematic changes to labor protection laws could contribute in the long-term to improving preparedness in emergency situations, as well as overall social and health equity.

As a social determinant of health, PSL has ramifications for one's health, well-being, and quality of life.[23,24] PSL makes an employee 60% more likely to receive an influenza vaccination and engage with medical and cancer screenings without forfeiting their income or jobs.[3] An additional study found that people without PSL were three times as likely to delay needed treatment due to concerns about the immediate costs of the treatment and related costs of wage loss. This relationship does not change when controlling for health status, education level, and income level.[25] The impact of PSL also applies to immediate family members, as parents who had PSL were more likely to take time off to care for children when needed. Furthermore, low-income children were less likely to have parents who had PSL.[26] The effects of this social determinant for an individual also extend to the community at large; one study estimated that due to a lack of PSL, 7 million people were additionally infected as a result of "presenteeism" in the workplace during the H1N1 pandemic.[27] A separate study estimated that Connecticut's PSL law resulted in a 14.8% reduction in the spread of illness in 2013.[6] Taken

Table 1
Multivariable mixed effects model: paid sick leave vs weekday workplace mobility.

Coefficient	β (95% CI)	*P* value[a]
Paid sick leave (reference: no)	−3.13 (−5.92, −0.34)	0.039
Yes		
Temporal components	−1.87 (−1.91, −1.82)	<0.001
Prepolicy effect	21.0 (5.64, 36.3)	0.053
Instantaneous effect	1.94 (1.89, 1.99)	<0.001
Postpolicy effect		
Health metrics		
New cases per 100,000	−0.03 (−0.04, −0.03)	<0.001
Travel metrics	−0.59 (−0.94, −0.24)	0.004
Average commute time (minutes)		
Average commute time on public transit (minutes)	−0.03 (−0.15, 0.09)	0.630
Economic metrics		
Unemployment rate (%)	−0.35 (−0.45, −0.26)	<0.001
2017 median household income ($10,000 USD)	0.19 (−0.91, 1.28)	0.742
Labour Overall Index Score	−0.03 (−0.08, 0.03)	0.339
MIT living wage (%)	0.36 (−0.75, 1.47)	0.534
Annual state GDP for 2019 (trillion USD)	−1.39 (−4.15, 1.37)	0.334
Poverty rate (%)	0.50 (0.07, 0.94)	0.035
Dominator labor sector (reference: trade, transportation, and utilities)	1.38 (−2.01, 4.77)	0.433
Education and health services	0.14 (−1.80, 2.07)	0.891
Government	2.20 (−3.68, 8.08)	0.471
Leisure and hospitality	7.34 (0.59, 14.1)	0.045
Manufacturing	1.01 (−4.47, 6.48)	0.722
Professional and business services		
Demographic metrics		
Black (%)	0.02 (−0.11, 0.14)	0.784
Hispanic (%)	−0.01 (−0.11, 0.10)	0.879
Asian (%)	0.01 (−0.30, 0.32)	0.933
Politics and policy		
Paid family leave (reference: no)	3.49 (−1.83, 8.81)	0.212
Yes		
Required pay reporting (reference: no)	0.22 (−4.93, 5.37)	0.934
Yes		
Split shift pay 2019 (reference: no)	−4.85 (−12.4, 2.74)	0.224
Yes		
Advanced shift notice 2019 (reference: no)	6.62 (−2.54, 15.8)	0.171
Yes		
Job-protected leave for non-FMLA workers 1 year on job (reference: no)	−1.20 (−4.37, 1.97)	0.466
Pregnant workers only	−3.47 (−7.15, 0.23)	0.080
Yes		
Job-protected leave longer than federal FMLA (reference: no)	1.23 (−1.96, 4.42)	0.458
Pregnant workers only	2.35 (−3.43, 8.13)	0.434
Yes		
Election results coding (reference: split)	−1.28 (−4.63, 2.07)	0.462
All democrat	−5.64 (−9.12, −2.17)	0.004
Mostly democrat	−1.06 (−4.52, 2.41)	0.556
Mostly republican	−0.81 (−3.40, 1.78)	0.545
All republican		
Other		
State population (1000 square miles)	−1.12 (−2.04, −0.20)	0.027

CI, confidence interval; GDP, gross domestic product; MIT, Massachusetts Institute of Technology; FMLA, Family Medical Leave Act.
[a] Values derived from a mixed-effects model with a nested random effect for state and date. The outcome of interest is percent change in weekday workplace mobility as determined from Google COVID-19 Community Mobility Reports.

together, these findings suggest that PSL plays a pivotal role in the well-being of both the individual with PSL, as well as their immediate colleagues and family.

Although the present study is the first to examine the impact of pre-existing state PSL on weekday workplace mobility during the COVID-19 pandemic, it has some limitations. First, publicly available covariate data were compiled across multiple sources and were measured at different points in time. Future work should attempt to standardize the time frame of analysis so that steps can be made toward establishing causality. Second, analysis was limited to the early stages of the COVID-19 pandemic, presenting future opportunities to examine the long-term impacts of pre-existing state PSL on workplace mobility. However, given the substantial drop in mobility that occurred in March 2020, it may be valuable for future work to explore this period in-depth. The substantial drop that occurs within this period is likely not associated with paid sick leave; rather, it corresponds to the mandatory stay-at-home orders, non-essential business closures, and declarations of emergencies that occurred within states during this period. We chose the date of FFCRA implementation (April 1) as our point of interest in part because it occurred after a majority of these state-level announcements took place, and we hypothesize that this may have biased our findings toward the null. Further quantification of the impact of stay-at-home orders and non-essential business closures on weekday workplace mobility is outside the scope of the present work.

Third, given the ecological nature of the study, future work is necessary to quantify the direct, person-level impact of pre-existing state PSL on workplace mobility. Fourth, Google COVID-19 Community Mobility Reports may not be representative of all populations (e.g. those without access to a cellular device). One limitation of these data is that they are not nationally representative, as there are discrepancies across age,

income bracket, and urban/rural divides for who owns a smartphone.[28] However, given that in recent decades, US public health policy has tilted toward states and that states have been at the forefront of the implementation of the American COVID-19 response, a state-by-state comparison of Google Mobility data allows for insight into each state's pandemic response and how it compares with others.[18,29,30] Because of the overwhelming heterogeneity of the United States, state-by-state observations are crucial to understanding the larger national picture. Fifth, the calculation of daily changes relative to a baseline in January and February 2020 (as opposed to a full year) may result in some seasonal biases. This may bias results away from the null, as individuals may be less likely to take off work during January and February compared with the following months. It should also be noted that states with and without pre-existing state PSL policies are spread across the United States. Per US Census Region, of the states without PSL, 31% are in the Midwest, 8% are in the Northeast, 38% are in the South, and 23% are in the West.[5,31] Of the states with PSL, 50% are in the Northeast, 17% are in the South, and 33% are in the West.[5,31] The geographic heterogeneity likely counteracts seasonal effects that may come from clusters of adjacent states. It is also important to note that the Google Mobility data analyzed were specifically with respect to how much time people spent in their workplace settings; depending on the type of work, this movement is expected to be less prone to seasonal influence than other types of movement (i.e. for recreation). Finally, this study is limited to PSL, and the evaluation of additional economic policies, such as medical leave for family members, flexible work hours, remote work policies, and flexibility in shift work, could offer more nuanced perspectives.

PSL is fundamental to preserving the health of the workforce, particularly during times of crisis. The results presented here suggest that pre-existing state policies may enhance the effectiveness of emergency legislation, although long-term, systemic labor protection laws remain crucial. Successful implementation of such laws requires an equity-based approach that considers addressing disparities in access to labor benefits, thoughtful outreach strategies through clear and consistent communication to all labor force members, and rigorous oversight and enforcement from state and federal labor departments and boards to both ensure compliance by employers and maximize the potential for success.[21]

Author statements

Acknowledgements

The authors would like to acknowledge Neha Shah, Tarek Eshak, Samia Baig, and Matthew Pietrosanu for their contributions to the initial stages of this work.

Ethical approval

None required.

Funding

None declared.

Competing interests

M.S.M. has previously received an unrestricted award from Google Research that is unrelated to the work presented in this paper. A.N.D. has received grant funding from the International Society for Infectious Diseases and PandemicTech and is a committee member on the Publications and Global Health Committees for IDSA.

Author contributions

FH, AD, SH, AND, and MSM conceptualized the work; FH, AD, and SH curated the data; all authors designed the analysis; CCP conducted the analysis; FH, CCP, AD, and SD wrote the original draft; all authors reviewed the final draft.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2022.08.019.

References

1. CDC. *COVID-19 and your health* [Internet]. Centers for Disease Control and Prevention; 2020 [cited 2021 Oct 28]. Available from: https://www.cdc.gov/coronavirus/2019-ncov/your-health/quarantine-isolation.html.
2. OECD. Paid sick leave to protect income, health and jobs through the COVID-19 crisis [Internet]. OECD. [cited 2021 Aug 1]. Available from: https://www.oecd.org/coronavirus/policy-responses/paid-sick-leave-to-protect-income-health-and-jobs-through-the-covid-19-crisis-a9e1a154/.
3. Lamsal R, Napit K, Rosen AB, Wilson FA. Paid sick leave and healthcare utilization in adults: a systematic review and meta-analysis. *Am J Prev Med* 2021;**60**(6):856–65. Jun 1.
4. Böckerman P, Kanninen O, Suoniemi I. A kink that makes you sick: the effect of sick pay on absence. *J Appl Econom* 2018;**33**(4):568–79.
5. KFF. *Paid family and sick leave in the U.S. [Internet]*. KFF; 2020 [cited 2021 Aug 22]. Available from: https://www.kff.org/womens-health-policy/fact-sheet/paid-family-leave-and-sick-days-in-the-u-s/.
6. Appelbaum E, Milkman R, Kroeger T. *Good for business? Connecticut's paid sick leave law*. 2014.
7. Paid sick leave in Washington state: evidence on employee outcomes, 2016–2018 | AJPH | Vol. 110 Issue 4 [Internet]. [cited 2022 Aug 9]. Available from: https://ajph.aphapublications.org/doi/10.2105/AJPH.2019.305481.
8. Xia J, Hayes J, Gault B, Nguyen H. *Paid sick days access and usage rates vary by race/ethnicity, occupation, and earnings [Internet]*. 2016. Feb [cited 2021 Aug 22]. Available from: https://iwpr.org/iwpr-publications/briefing-paper/paid-sick-days-access-and-usage-rates-vary-by-race-ethnicity-occupation-and-earnings/.
9. Families first Coronavirus response Act: employee paid leave rights | U.S. Department of labor [Internet]. [cited 2021 Aug 22]. Available from: https://www.dol.gov/agencies/whd/pandemic/ffcra-employee-paid-leave.
10. Glynn SJ. *Coronavirus paid leave exemptions exclude millions of workers from coverage [Internet]*. Center for American Progress; 2020 [cited 2021 Aug 1]. Available from: https://www.americanprogress.org/issues/economy/news/2020/04/17/483287/coronavirus-paid-leave-exemptions-exclude-millions-workers-coverage/.
11. Pichler S, Wen K, Ziebarth NR. COVID-19 emergency sick leave has helped flatten the curve in the United States. *Health Aff (Millwood)* 2020;**39**(12):2197–204. Dec 1.
12. Andersen M, Maclean JC, Pesko MF, Simon KI. *Paid sick-leave and physical mobility: evidence from the United States during a pandemic* [Internet]. National Bureau of Economic Research; 2020. May [cited 2021 Aug 1]. (Working Paper Series). Report No.: 27138. Available from: https://www.nber.org/papers/w27138.
13. COVID-19 community mobility report [Internet]. COVID-19 Community Mobility Report. [cited 2021 Aug 22]. Available from: https://www.google.com/covid19/mobility?hl=en.
14. MAP: state labor index (Best states to work) [Internet]. [cited 2021 Sep 15]. Available from: https://www.oxfamamerica.org/explore/countries/united-states/poverty-in-the-us/best-states-to-work/.
15. Times TNY. Coronavirus in the U.S.: latest map and case count. The New York Times [Internet]. 2020 Mar 3 [cited 2021 Sep 15]; Available from: https://www.nytimes.com/interactive/2021/us/covid-cases.html.
16. Bureau UC. American Community Survey (ACS) [Internet]. The United States Census Bureau. [cited 2021 Sep 15]. Available from: https://www.census.gov/programs-surveys/acs.
17. Historical Election Results | USAGov [Internet]. [cited 2021 Sep 15]. Available from: https://www.usa.gov/election-results.
18. Warner ME, Zhang X. Social safety nets and COVID-19 stay home orders across US states: a comparative policy analysis. *J Comp Policy Anal Res Pract* 2021;**23**(2):176–90. Mar 4.
19. Selden TM, Berdahl TA. Risk of severe COVID-19 among workers and their household members. *JAMA Intern Med* 2021;**181**(1):120–2. Jan 1.
20. Sze S, Pan D, Nevill CR, Gray LJ, Martin CA, Nazareth J, et al. Ethnicity and clinical outcomes in COVID-19: a systematic review and meta analysis. *EClinicalMedicine* 2020;**29**:1–17.

21. Rauhaus B, Johnson A. Social inequities highlighted by the prolonged pandemic: expanding sick leave. *J Public Nonprofit Aff* 2021;**7**(1):154–63. Apr 1.
22. Palme M, Persson M. Sick pay insurance and sickness absence: some European cross-country observations and a review of previous Research. *J Econ Surv* 2020;**34**(1):85–108.
23. Social determinants of health | healthy people 2020 [Internet]. [cited 2021 Dec 3]. Available from: https://www.healthypeople.gov/2020/topics-objectives/topic/social-determinants-health/interventions-resources.
24. Romich JL. Local mandate improves equity of paid sick leave coverage: seattle's experience. *BMC Publ Health* 2017;**17**(1):60. Dec.
25. DeRigne L, Stoddard-Dare P, Quinn L. Workers without paid sick leave less likely to take time off for illness or injury compared to those with paid sick leave. *Health Aff (Millwood).* 2016;**35**(3):520–7. Mar 1.
26. Clemans-Cope L, Perry CD, Kenney GM, Pelletier JE, Pantell MS. Access to and use of paid sick leave among low-income Families with children. *Pediatrics* 2008;**122**(2):e480–6. Aug 1.
27. Drago R, Miller K. *Sick at work: infected employees in the workplace during the H1N1 pandemic [Internet]*. IWPR; 2010 [cited 2021 Nov 23]. Available from: https://iwpr.org/iwpr-general/sick-at-work-infected-employees-in-the-workplace-during-the-h1n1-pandemic/.
28. Pew Research Center. Mobile fact sheet [Internet]. Pew Research Center: Internet, Science & Tech. [cited 2022 Aug 9]. Available from: https://www.pewresearch.org/internet/fact-sheet/mobile/.
29. Turnock BJ, Atchison C. Governmental public health in the United States: the implications of federalism. *Health Aff (Millwood)* 2002;**21**(6):68–78. Nov.
30. Huberfeld N, Gordon SH, Jones DK. Federalism complicates the response to the COVID-19 health and economic crisis: what can be done? *J Health Polit Policy Law* 2020;**45**(6):951–65. Dec 1.
31. Bureau UC. Geographic Levels [Internet]. Census.gov. [cited 2022 Aug 9]. Available from: https://www.census.gov/programs-surveys/economic-census/guidance-geographies/levels.html.

Public Health 215 (2023) 124–130

Contents lists available at ScienceDirect

Public Health

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

Original Research

What is the evidence that advertising policies could have an impact on gambling-related harms? A systematic umbrella review of the literature

E. McGrane [a], H. Wardle [b], M. Clowes [a], L. Blank [a], R. Pryce [a], M. Field [c], C. Sharpe [d], E. Goyder [a, *]

[a] School of Health & Related Research (ScHARR), University of Sheffield, Sheffield, UK
[b] School of Social and Political Sciences, University of Glasgow, Glasgow, UK
[c] Department of Psychology, University of Sheffield, Sheffield, UK
[d] School of Public Health, Imperial College London, UK

ARTICLE INFO

Article history:
Received 27 September 2022
Received in revised form
25 November 2022
Accepted 25 November 2022

Keywords:
Gambling-related harm
Advertising
Inequalities
Health policy

ABSTRACT

Objective: To summarise the evidence on the impacts of gambling-related advertising that could lead to gambling-related harm, including impacts on vulnerable individuals and inequalities in the distribution of harms.
Study design: An umbrella review of studies investigating the impact of gambling advertising.
Methods: A review was undertaken of systematic reviews of qualitative, quantitative and mixed method studies reporting outcomes associated with gambling advertising and marketing. The search strategy included database searches (Web of Science, PsycInfo) and website searches. The quality of the included reviews was determined using A MeaSurement Tool to Assess systematic Reviews 2.
Results: 1024 papers were identified by database searches. Eight systematic reviews, including 74 unique studies, met inclusion criteria. Included studies, using quantitative and qualitative methods, consistently support the existence of a causal relationship between exposure to advertising of gambling products/brands and more positive attitudes to gambling, greater intentions to gamble and increased gambling activity at both individual and population level. There is evidence of a 'dose–response' effect; greater advertising exposure increases participation which leads to a greater risk of harm. There was more evidence for the impact on children and young people and for those already at risk from current gambling activity with those most vulnerable more likely to be influenced.
Conclusion: Gambling advertising restrictions could reduce overall harm and mitigate the impact of advertising on gambling-related inequalities. Public health harm prevention strategies should include policies which limit exposure to advertising, particularly among children and vulnerable groups.

Introduction

There is substantial international and UK-specific evidence base on the range of harms related to gambling[1] and the risk factors that predict an increased risk of harm from gambling.[2] Causal pathways between gambling, health and wellbeing at both individual and population levels include the effects of financial loss and debt, as well as wider impacts on relationships, education, employment and crime. Recent national and regional data have quantified the scale of the associated harms and economic costs for the UK.[3,4]

Whilst gambling disorder may be the most serious and widely recognised gambling-related health condition, harm to health may occur even at relatively low levels of gambling activity. The various harms to health and wellbeing may be the result of diverse mechanisms including impacts on mental health, relationships and financial stresses.[5] Risks of gambling-related harms are associated with a range of individual, psychosocial, political, economic and commercial factors. Thus gambling-related harms represent a significant potential driver of health inequalities because those already experiencing financial, social and mental health

* Corresponding author. ScHARR, University of Sheffield, Regent Court, 30 Regent St, Sheffield, S1 4DA, UK. Tel.: +114 222 0783.
E-mail address: e.goyder@sheffield.ac.uk (E. Goyder).

https://doi.org/10.1016/j.puhe.2022.11.019

disadvantage are also at increased risk of experiencing gambling-related harm (the 'gambling harm paradox').[6]

As one aspect of industry marketing strategies, gambling advertising is ubiquitous across a wide range of media including both more traditional forms (e.g. television, newspapers, outdoor and point of sale advertising) and more recent channels (e.g. internet and social media advertising). Recent policy reports and evidence reviews suggest that direct evidence of a causal relationship between exposure to gambling advertising and gambling-related harms is not easily obtainable. Yet a wealth of indirect evidence exists showing an association between advertising and attitudes and behaviour as well as an association between attitudes and behaviour and subsequent risk of harms.

We undertook a synthesis of review evidence on the relationship between advertising and attitudes, intentions and behaviours which, in turn, may be associated with an increased risk of gambling-related harms. We aimed to summarise evidence exploring the relationship between exposure to advertising and attitudes and behaviour that can lead to harms and to explore evidence on the impact of advertising for individuals and communities known to be more vulnerable to gambling-related harms.

Methods

Given the large volume and diversity of research on gambling adverting and its effects, we undertook an umbrella review of relevant systematic reviews which included primary studies of the impact of gambling advertising.

Search strategy and selection criteria

The search, which comprised subject headings and free-text terms, was initially developed and run on PsycINFO before being adapted for Web Of Science (Science Citation Index and Social Science Citation Index) (see Supplementary File 1 for search strategy). Database searches, undertaken in February 2022 and citation searches in March 2022, were limited to English language reviews published since 2000. Database searching was accompanied by scrutiny of reference lists and citations of included papers, searches for grey literature including a search of relevant key websites (see Supplementary File 1) in March 2022. Two reviewers (EM and EG) independently undertook study selection. Uncertainties on study inclusion were resolved by discussion between the two reviewers and among the wider review team as required. Inclusion criteria were specified as:

Population: Any population/region exposed to gambling advertising including subgroups e.g. children and young people; groups at higher risk of gambling-related harms; those already experiencing gambling-related harms and/or seeking treatment.

Exposure: This included exposure to any form of gambling advertising including experimentally-manipulated or observed exposure, or self-reported recall of exposure. All forms of advertising were included (broadcast and print media, outdoors, on line and point of sale advertising) but other marketing strategies (e.g. sponsorship of events or charitable funding) which did not use advertising to raise awareness or encourage product use were outside the scope of this review.

Outcomes: Any outcome related to gambling in terms of attitudes, intentions or behaviour.

Studies: We included all reviews which described a systematic method for identifying included evidence, reviews including both published and unpublished ('grey') literature. We excluded reviews which only covered advertising or marketing content or strategies and did not include studies related to the impact of advertising.

Data analysis

Data extraction was performed by one reviewer (EM) and checked for accuracy and consistency by a second (EG). The data extraction focused on identifying the main associations between exposure to advertising, attitudes, intentions and behaviour that were reported and the number and nature of the primary studies on which the findings were based. The quality of the included reviews was determined by two reviewers (EG and EM) independently using A MeaSurement Tool to Assess systematic Reviews 2 (see Supplementary File 3 for findings).[7] The extracted data was synthesised narratively due to the diverse nature of the evidence.

Results

After de-duplication, the initial database searches generated 1024 records, of which 24 were retrieved as full papers. Five of these met our inclusion criteria (see Supplementary File 2 list of excluded reviews). Two additional reviews were identified from website searches and one additional review was identified by a topic expert. No additional reviews were identified by reference or citation searches (Fig. 1 shows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram).

In total, eight systematic reviews met inclusion criteria. The scope and main findings of individual reviews are summarised in Table 1 (see Supplementary File 3 for quality appraisal findings). Three reviews included both studies of those already at risk from their gambling and general populations (both adults and children).[8–10] Two reviews only included studies of children[11,12] and one review focused on migrant communities.[13] Two reviews commissioned to inform policymaking focused on intervention policies.[14,15] One review, funded by a national charity that receives funds from the gambling industry (GambleAware), explicitly stated a funding source.[9]

Bouguettaya (2020)[8] was the only review to include a quantitative synthesis of the relationship between exposure to gambling advertising and gambling attitudes, intentions and behaviours. In total, 28 papers published between 2000 and 2019 were included in this study (24 in the meta-analysis). All included studies measured the impact of gambling advertising, six qualitative, 20 quantitative and two mixed methods. The correlation coefficients on attitude ranged from $r = 0.12$ to $r = 0.62$ (mean $r = 0.40$), those on intentions ranged from $r = 0$ to $r = 0.2$ (mean $r = 0.05$), those on behaviour ranged from $r = -0.8$ to $r = 0.68$ (mean $r = 0.24$). The aggregated meta-analysis correlation coefficients were positive but not statistically significant for intentions and attitudes. The aggregated coefficient on the behavioural variable was positive and statistically significant; this category included a much larger number of papers with higher numbers of significant effect sizes. Thus, the quantitative evidence is strongest for the relationship between advertising exposure and behaviour. Longitudinal Ecological Momentary Assessment studies which asked individuals to report their actual exposure to direct messaging as it occurred reported the highest number of significant effects.

Cross-sectional evidence on children and adolescents showed that higher exposure to advertising is associated with higher gambling rates and severity. Some studies found links with gambling intent amongst adolescents and other studies found links with attitudes. Cross sectional evidence reports that higher exposure to advertising is particularly associated with stronger intentions and influences betting behaviour in those who are current and higher risk gamblers. Higher risk gamblers also have a higher self-reported exposure to advertising and tend to hold more positive attitudes about advertising. In general, qualitative findings indicate that high levels of exposure to advertising normalises

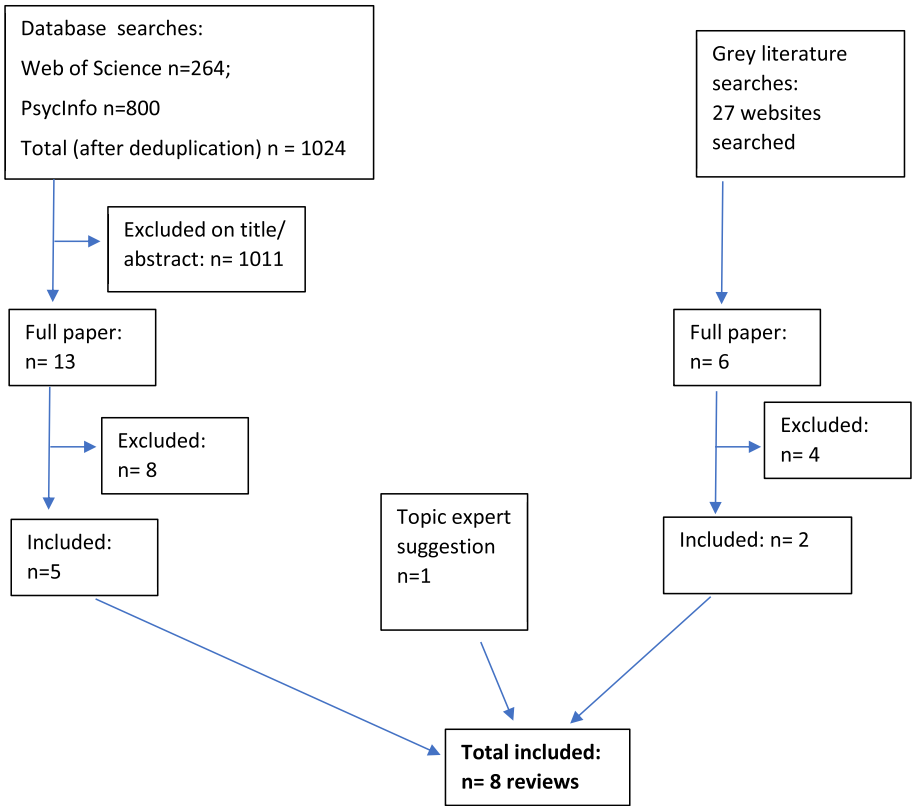


Fig. 1. PRISMA Flow Diagram of studies.

gambling, creates positive attitudes and acts as an incentive to gamble.

Included studies suggested that advertising would not encourage young people to start gambling, but instead generate difficulties for existing young gamblers, especially those experiencing problems, who reported that gambling advertising had some impact on their behaviour. Additionally, advertising was observed to have a priming effect, teaching children how to place a bet.

Newall et al. (2019)[9] carried out a systematic review of gambling marketing research published between 2014 and 2018. Of the 46 papers included, 27 focused on advertising and behaviour. Seven were unique to this review. Of these 27 behavioural papers, 18 included an explicit discussion of the impact of advertising exposure, ten quantitative, six qualitative, two mixed-methods. Longitudinal observational studies also found that advertising exposure was associated with increased gambling expenditure. An experimental study found that 'push' notifications resulted in larger and riskier bets being placed. The cross-sectional quantitative evidence showed that higher risk gamblers have a greater awareness of and exposure to gambling advertising. They are also more likely to report that it has increased their involvement in gambling, notably in studies looking at sports betting.

Guillou-Landreat (2021)[10] identified 21 studies on the digital marketing, of which nine specifically discuss the impact of advertising: five quantitative, three qualitative and one mixed-method. Longitudinal studies suggest that exposure to different forms of advertising is consistently associated with either an increased probability of betting, increased expenditure on betting or increased intention to bet amongst sports bettors. Cross-sectional evidence demonstrated a dose-response effect with the average number of inducement offers received significantly predicting the number of unplanned bets placed before and during sports matches. Cross-sectional evidence showed the perceived self-reported impact of advertising on behaviour is a significant predictor of problematic gambling severity. A higher percentage of those experiencing moderate risk or problem gambling reported that social media promotions for gambling increased their problems compared with low-risk and 'non-problem' gamblers. These promotions also increased impulsive betting for higher-risk gamblers.

Labrador et al. (2021)[12] summarised the last 10 years of literature on gambling advertising to adolescent and youth populations. Of 31 included studies, 17 studies specifically discussed the impact of advertising, seven quantitative, eight qualitative and two mixed methods. All included studies were cross-sectional and descriptive. In these studies, most adolescents and youth report that advertising would not influence their own behaviour, and only a small proportion also said that they intended to gamble at age 18. However, adolescents who have already engaged in some form of gambling (the majority in most studies) have enhanced recall of advertising. They were familiar with the content of gambling advertising and some believed that it misinforms people, pushing them to gamble. Some studies found that exposure to advertisements significantly predicted adolescent gambling behaviour, but only for people over the age of 18. Despite these discrepancies, advertising including pricing promotions and those which appear visually appealing were consistently mentioned as being the most effective strategy for motivating young people to participate in gambling activities. Bouguettaya et al. reported similar findings.[8] A large percentage of adolescents experiencing problem gambling reacted to advertising with a want to engage in betting. They also reported an oversaturation of advertising and marketing, contributing to the normalisation of gambling.

Table 1
Characteristics of included reviews.

Authors	Title	No. of relevant studies included	Exposure variables identified	Outcome variables identified	Subgroups and modifying factors identified	Limitations and potential sources of bias identified within included primary studies	Summary of results
Bouguettaya et al. (2020)[8]	The relationship between gambling advertising and gambling attitudes, intentions and behaviours: a critical and meta -analytic review	27 (11 -attitudes 8—intentions 2—-behaviour)	Advertising observed (recall), fake or real advertisements, ban on EGMs, expenditure on advertising, reporting watching a show with embedded advertising, online and offline advertising	Intent to gamble (e.g. likelihood of placing a bet), problem gambling, actual betting, past gambling, attitudes (e.g. feelings)	Children and young people; 'problem' and 'non-problem' gamblers	Lack of high quality research. Most rely on recall or self-report which risks reporting bias. Quantitative papers suffer from poor methodological and statistical reporting. Reverse causation cannot be ruled out due to lack of longitudinal and experimental studies. Majority evidence from Australia (16 of 27)	Attitudes and intentions: exposure likely to be associated with more positive attitudes and greater intentions to gamble. Behaviour: exposure likely to increase gambling and problem gambling behaviour.
Newall et al. (2019)[9]	Gambling marketing from 2014 to 2018: A literature review	19—perception 8—behaviour	Different types of advertising e.g. free bets/'risk-free'/sports related/ casino games	Perceptions related to recall; awareness; normalisation; understanding; susceptibility	Children; 'problem' and 'non-problem' gamblers	Largely retrospective and recall of advertising and of behaviour both subject to recall/reporting bias. Majority of evidence from Australia; little from other countries	Perception: more negative for active gamblers; children may be influenced/ misled Behaviour: exposure prompts more frequent and riskier gambling
Guillou-Landreat et al. (2021)[10]	Gambling Marketing Strategies and the Internet: What Do We Know? A Systematic Review.	21 (9—behaviour 12— content/ perceptions)	Self-reported exposure to advertising, uptake of inducements, receiving direct messages, number of gambling accounts	Subjective (self-reported) influence on betting, impact of specific features of advertising	Children and young people; 'problem' and 'non-problem' gamblers	Potential cultural bias as most studies are from New Zealand and Australia. Potential selection bias as it did not include studies on traditional media (i.e. TV, radio, press). Both limit the generalisability of results. Limited discussion of the methodological limitations of the literature (relying on self-report, lack of causal evidence)	Behaviour: increased accessibility and use of promotions as influencing behaviour. Problem of impulse sports betting, especially for problem/at-risk gamblers.
Wardle (2019)[11]	Perceptions, people and place: Findings from a rapid review of qualitative research on youth gambling	21 (8—advertising and behaviour, 13—other)	N/A	Self-reported opinions on gambling advertising (focus groups/interviews)		Methodologies (including sampling) tend to be poorly reported, and there is a lack of detailed analysis in some papers. Some studies had wide ranging aims and lacked depth. Lack of evidence on gender and cultural differences (and socioeconomic). Potential for systematic biases in those who took part in the research due to this lack of diversity.	Perceptions: advertising seen to normalise gambling (especially in sport) Behaviour: bonus offers (free bets) as being the greatest incentive, advertising making young people 'want to bet'
Labrador et al. (2021)[12]	Exposure of adolescents and youth to gambling advertising: A systematic review	31 (23—behaviour/ attitude, 8—content)	Recall of brands/adverts, self-reported exposure to gambling advertising	Self-reported (perceived) impact of advertising	Gender as modifying factor (young people only included)	Some studies are older and might be outdated Papers rely on self-report which risks reporting and recall bias Cultural bias as most studies from Australia/Canada	Behaviour: most young people feel that advertising would not influence their behaviour. Some feel that it makes them want to bet. Promotions are seen as the most effective strategy to motivate participation. Attitudes: advertising normalises gambling, shows you how to do it Gender: some evidence that young males are more affected by advertising
Wardle et al. (2019)[13]	What do we know about gambling-related harm affecting migrants and migrant communities? A rapid review	38 (4—advertising and behaviour)	N/A	N/A	Migrant groups (not born in the country in the study)	Most studies are New Zealand/Australia so there may be a cultural bias	Exposure: might have increased exposure compared to their home jurisdictions; advertising targeted towards specific ethnic groups (e.g. built on significant cultural events) Behaviour: migrants tend to gamble less overall

(*continued on next page*)

Table 1 (*continued*)

Authors	Title	No. of relevant studies included	Exposure variables identified	Outcome variables identified	Subgroups and modifying factors identified	Limitations and potential sources of bias identified within included primary studies	Summary of results
Livingstone et al. (2019)[14]	Identifying effective policy interventions to prevent gambling-related harm	100 (4—advertising and behaviour)		Attitudes (whether advertising resonates with respondents)	Adolescents and young people	Overall study quality is weak (lack of large-scale studies, RCTs not feasible, significant influence from industry) Undeclared, incomplete or inaccurate declarations of COI in relevant sections of publications (some journals do not consistently enforce ISAJE guidelines) Lack of neutral terminology (e.g. "gaming" instead of "gambling", "play" instead of "use")	Behaviour: there is little evidence on the efficacy/effectiveness of restricting advertising: sports betting advertising resonates with younger fans (<18yrs) by normalising/legitimising it Attitudes: studies reported no effects from intervention or did not recommend school-based educational programmes.
Rodda (2020)[15]	A Rapid Review and Research Gap Analysis: A 2020 update.	215 (20—policy, {of which 8—advertising and behaviour) 195—other)		Attitudes, intentions (intended betting), and behaviour (actual betting)	Adolescents and young people; 'problem' and 'non-problem' gamblers	Empirical literature is mostly exploratory and cross-sectional (low quality). Policy literature (section including the advertising literature) assigned a particularly low quality rating: 75% low/very low, 10% high quality	Behaviour: included studies consistently report an association between advertising, incentives, and unwanted gambling behaviours (e.g. 'problem' gambling) Evidence is well established (particularly for in-game promotions)

Wardle (2019)[11] undertook a rapid systematic review of qualitative literature on the perceptions, determinants and gambling experiences of young people to understand the impact of gambling and the precursors of gambling behaviour in the future. 21 papers were included, of which seven discussed the impact of advertising five qualitative, two mixed-methods. The mixed methods research shows that children and young people have high awareness and recall and therefore exposure to gambling advertising. When asked directly, young people report that advertising does not impact their behaviour. Despite this, the anecdotal evidence in qualitative studies suggests that children and young people are influenced by gambling advertising. Evidence shows that young people were influenced most by promotions, such as bonus bets. Many view these incentives as 'free money', encouraging them to sign-up with multiple operators. Incentives were seen to 'lure' young people into gambling with the promise of winning or the 'fear of missing out'. Targeted and personalised advertising was also recognised as gambling companies 'pursuing' young people. Generally, advertising is seen to alter the gambling environment for young people, normalising the activity and making young people feel a need to bet.

Another rapid systematic review by Wardle et al. (2019)[13] explored gambling participation, motivations, harms and provision of support for migrant populations. Of 38 included studies, three papers published between 2009 and 2016 discussed the potential impact of advertising two qualitative and one literature review. The literature suggests that advertising may be specifically targeted towards migrant groups to encourage participation in gambling. Some advertising campaigns have been reported to utilise significant cultural events to attract ethnic minorities to gambling activities, using people as cultural symbols.

A systematic review by Livingstone et al. (2019)[14] sought to critically assess the literature on harm minimisation related to Electronic Gaming Machines (EGM) and online betting. Out of 100 articles related to gambling, four studies looked at gambling advertising. Of these, three were unique to this review and two looked specifically at the impact of advertising; the widely cited Binde (2014)[16] literature review and a qualitative study found that children have high recall of gambling advertising and brands. Children and young people were most aware of advertising linked with sports, which is seen to normalise gambling. Some children reported wanting to bet on sports due to the widespread advertising of sports gambling.

Rodda (2020)[15] undertook a rapid systematic review of the gambling literature with a focus on harm minimisation. This review included 215 studies covering seven research questions. 20 papers were relevant to the research question about policy and 14 of these included a discussion on gambling advertising. This search captured four systematic reviews already included in this umbrella review and an additional five unique primary papers. The longitudinal and cross-sectional evidence shows that advertisements for gambling are linked to a greater likelihood of betting, intention to bet, and expenditure on betting. Longitudinal evidence suggests that advertising influences the frequency and size of bets amongst existing bettors, but these results do not vary by gambling risk level. Cross-sectional evidence suggests that young people have high recall of gambling advertising. Experimental evidence indicates that higher risk gamblers experience higher physiological desire when viewing advertising. They also have higher overall desire ratings for advertisements and subsequently higher rates of gambling harm.

Overall, these reviews consistently reported that exposure to advertising is associated with more positive attitudes and greater reported intentions to gamble. They find a direct association between exposure to advertising and gambling activity, with a 'dose

response' effect whereby greater exposure to or awareness of advertising is associated with more gambling activity and higher risk gambling activity. Associations between exposure to advertising, positive attitudes to gambling and more risky gambling behaviour are generally found to be greater for those individuals who are already at risk of harms and higher levels of exposure to advertising are directly related to gambling severity scores (as measured by the Problem Gambling Severity Index).[16] Children and young people consistently report a high level of awareness of and exposure to advertisements with both parents and children reporting that advertising has normalised gambling as a risk-free leisure activity.

Discussion

The number of very recent systematic reviews included in this umbrella review reflects a significant increase in both primary research and evidence synthesis in the field of gambling advertising. A particularly large number of relevant primary studies and reviews have been published in the last three years. Limited time and resources meant that our review could not be fully comprehensive. In particular, we were not able to include reviews published on languages other than English that might have expanded the coverage of regions with different policy approaches to gambling advertising.

A traditional systematic review methodology synthesises findings from similar studies, addressing the same research question and generally using the same or similar methods. In contrast, our review aimed to bring together the diversity of approaches and evidence relevant to understanding the causal pathway between advertising and gambling-related harms. Our review thus included a broad range of evidence, foregoing the ability to undertake a quantitative synthesis and identification of the consistency of evidence at the primary study level. In this context, it is notable that there was strong consensus across reviews in terms of conclusions despite widely differing synthesis methods. Using these methods, we were still able to identify a large and diverse body of evidence on the relationship between advertising and marketing related to gambling activities, products and brands and a wide range of outcomes.

Different study designs provided different types of evidence examining the relationship between advertising and harms. Quasi-experimental studies and surveys have provided consistent evidence for an association between exposure to advertising and gambling-related outcomes and increasing evidence of a 'dose–response' relationship, greater exposure being associated with larger effect sizes. The experimental and qualitative studies provide detailed evidence regarding causal mechanisms. Experimental studies reveal the role of intentions: advertising exposure directly influences decisions to gamble and to participate in more risky gambling. Qualitative studies based on focus groups and in depth interviews explored how, and why, some subgroups may be particularly susceptible to harmful responses to advertising. They demonstrate how social effects of advertising, such as 'normalisation', may lead to harm. All study designs contribute to the evidence for a dose–response relationship whereby increasing exposure has an increasing impact. Similarly, all study types provided evidence specific to the impact of advertising on vulnerable groups who may be at a higher risk of harm from advertising exposure.

Several widely cited reviews, that did not meet our inclusion criteria, support our overall findings of consistent associations between exposure to advertising and attitudes, intentions and gambling behaviour and that the relationship is strongest among those already at risk of harm from their gambling activity.[17–24] More recent primary studies provide additional support for consistent associations between exposure to advertising and gambling-related attitudes and behaviour, including higher risks of harmful gambling activities for children, young people and those already at risk of harm from their gambling behaviour.[25–32]

The evidence base does have significant limitations and is largely characterised by cross-sectional surveys and qualitative studies of self-reported exposure, attitudes and behaviour in the general population and experimental or quasi-experimental studies conducted with those already identified as at increased risk due to their gambling activity or seeking treatment. There is a notable lack of longitudinal studies. There are also gaps in relation to evidence related to some specific forms of advertising, particularly outdoor and point of sale advertising which are environmental exposures over which the individual has very little control. Given the dependence on self-report of gambling activity in this field, there is an urgent need of research to use more objective measures of both exposures to advertising and gambling activity (e.g. account data). However, it is also crucial that research in this field is independent of any risk of industry influence and of direct or indirect industry funding and that funding of research and authors' potential conflicts of interest are always comprehensively and transparently reported.

In the absence of definitive controlled studies, the substantial and consistent evidence base supports restrictions to reduce exposure to gambling advertising. This is particularly likely to reduce risk of harm to children and young people and among adults who are already vulnerable to, or experiencing, gambling-related harms. Such restrictions could not only reduce overall harm but also mitigate the impact of advertising on gambling-related inequalities. Public health harm prevention strategies should therefore include policies which limit exposure to advertising, particularly among children and vulnerable groups.

Gambling advertising restrictions could reduce overall harm and mitigate the impact of advertising on gambling-related inequalities. Public health harm prevention strategies should ideally include a range of policies which limit exposure to advertising, particularly among children and vulnerable groups. Policy evaluations of such restrictions, using methods that have already been successful in evaluating the impact of other advertising restrictions,[33–35] could also add significantly to the evidence base to inform future public health policy.

Author statements

Acknowledgements

This report was funded by the Greater London Authority to provide an independent evidence briefing on the evidence base related to the health and inequalities impacts of gambling advertising. The views expressed are those of the authors and not necessarily those of the GLA or members of the project reference group. The reference group provided advice and expert guidance to the project and provided feedback on a draft of this report.

Ethical approval

None sought.

Funding

None declared.

Competing interests

HW is the Deputy Chair of Advisory Board for Safer Gambling and has had funding from the Gambling Commission and the Gambling Research Exchange Ontario, and travel funded by the Alberta Gambling Research Institute and Cyprus Betting and Gaming Commission. We declare no other competing interests.

Contributors

EM, EG, HW, MC and CS drafted the study protocol. MC developed the search strategy and undertook the literature searches, and EM, EG and HW contributed to the screening process and selection of included studies. EM and EG extracted data and completed the quality assessment independently. All authors (EM, HW, MC, LB, RP, MF, CS, EG) contributed to the data synthesis and interpretation of findings. All authors (EM, HW, MC, LB, RP, MF, CS, EG) critically reviewed and approved the manuscript.

Data sharing

All the data included in this review are in the public domain in the form of journal articles and/or reports available on line or from the corresponding author.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2022.11.019.

References

1. Langham E, Thorne H, Browne M, Donaldson P, Rose J, Rockloff M. Understanding gambling related harm: a proposed definition, conceptual framework, and taxonomy of harms. *BMC Publ Health* 2016;**16**.
2. Public Health England. *Risk factors for gambling and harmful gambling: an umbrella review - a review of systematic reviews and metaanalyses*. London: PHE; 2021.
3. Public Health England. *Gambling-related harms evidence review: quantitative analysis of gambling involvement and gambling-related harms among the general population in England*. London: PHE; 2021.
4. Public Health England. *Gambling-related harms evidence review: the economic and social cost of harms*. London: PHE; 2021.
5. Wardle H, Reith G, Langham E, Rogers RD. Gambling and public health: we need policy action to prevent harm. *BMJ* 2019;**365**.
6. Raybould JN, Larkin M, Tunney RJ. Is there a health inequality in gambling related harms? A systematic review. *BMC Publ Health* 2021;**21**:305.
7. Biondi-Zoccai G. *Umbrella reviews. Evidence synthesis with overviews of reviews and meta-epidemiologic studies*. Cham, Switzerland: Springer International; 2016.
8. Bouguettaya A, Lynott D, Carter A, Zerhouni O, Meyer S, Ladegaard I, et al. The relationship between gambling advertising and gambling attitudes, intentions and behaviours: a critical and meta-analytic review. *Special Issue: Emotion, Motivation, Personality and Social Sciences: Gambling* 2020;**31**:89–101.
9. Newall PW, Moodie C, Reith G, Stead M, Critchlow N, Morgan A, et al. Gambling marketing from 2014 to 2018: a literature review. *Current Addiction Reports* 2019;**6**:49–56.
10. Guillou-Landreat M, Gallopel-Morvan K, Lever D, Le Goff D, Le Reste J-Y. Gambling marketing strategies and the internet: what do we know? A systematic review. *Front Psychiatr* 2021;**12**.
11. Wardle H. Perceptions, people and place: findings from a rapid review of qualitative research on youth gambling. *Addict Behav* 2019;**90**:99–106.
12. Javier Labrador F, Jose Estupina F, Vallejo-Achon M, Sanchez-Iglesias I, Gonzalez-Alvarez M, Fernandez-Arias I, et al. Exposure of adolescents and youth to gambling advertising: a systematic review. *An Psicolog* 2021;**37**:149–60.
13. Wardle H, Bramley S, Norrie C, Manthorpe J. What do we know about gambling-related harm affecting migrants and migrant communities? A rapid review. *Addict Behav* 2019;**93**:180–93.
14. Livingstone C, Rintoul A, de Lacy-Vawdon C, Borland R, Dietze P, Jenkinson R, et al. *Identifying effective policy interventions to prevent gambling-related harm*. Melbourne: Victorian Responsible Gambling Foundation; 2019.
15. Rodda S. *A rapid review and research gap analysis: a 2020 update*. NSW Responsible Gambling Fund; 2020.
16. Ferris J, Wynne H. *The Canadian problem gambling Index: final report*. 2021. Ottawa, Canada.
17. Binde P. *Gambling advertising: a critical research review*. London: Responsible Gambling Trust; 2014.
18. Griffiths MD. Does gambling advertising contribute to problem gambling? *Int J Ment Health Addiction* 2005;**3**:15–25.
19. Buchanan J, Shen Y. Gambling and marketing: a systematic literature review using HistCite. *Account Finance* 2021;**61**:2837–51.
20. Hörnle J, Carran MA. A sieve that does hold a little water – gambling advertising and protection of the vulnerable in the UK. *Leg Stud* 2018;**38**:529–48.
21. Lawn S, Oster C, Riley B, Smith D, Baigent M, Rahamathulla M. A literature review and gap analysis of emerging technologies and new trends in gambling. *Int J Environ Res Publ Health* 2020;**17**.
22. Lopez-Gonzalez H, Estévez A, Griffiths MD. Marketing and advertising online sports betting: a problem gambling perspective. *J Sport Soc Issues* 2017;**41**:256–72.
23. Parke A, Harris A, Parke J, Rigbye J, Blaszczynski A. Responsible marketing and advertising in gambling: a critical review. *The Journal of Gambling Business and Economics* 2015;**8**:21–35.
24. Planzer S, Wardle H. What we know about the comparative effectiveness of gambling regulation. *Eur J Risk Regulat* 2012;**3**:410–6.
25. Grant JE, Chamberlain SR, Schreiber LR, Odlaug BL. Gender-related clinical and neurocognitive differences in individuals seeking treatment for pathological gambling. *J Psychiatr Res* 2012;**46**:1206–11.
26. Kristiansen S, Severin-Nielsen MK. Adolescent gambling advertising awareness: a national survey. *Int J Soc Welfare* 2022;**31**:263–73.
27. Browne M, Bryden GM, Russell AM, Rockloff M, Hing N, Dittman C, et al. *Survey of South Australian adults'(18+) attitudes, beliefs and behaviours regarding sports betting Victoria: South Australia office for problem gambling*. 2021.
28. Syvertsen A, Erevik EK, Hanss D, Mentzoni RA, Pallesen S. Relationships between exposure to different gambling advertising types, advertising impact and problem gambling. *J Gambl Stud* 2021:1–18.
29. Syvertsen A, Pallesen S, Erevik EK, Mentzoni RA. Direct marketing experiences among individuals with current and lifetime gambling disorder. *Front Psychol* 2020;**11**.
30. Lopez-Gonzalez H, Griffiths MD, Jimenez-Murcia S, Estevez A. The perceived influence of sports betting marketing techniques on disordered gamblers in treatment. *Eur Sport Manag Q* 2020;**20**:421–39.
31. Wardle H, Critchlow N, Brown A, Donnachie C, Kolesnikov A, Hunt K. The association between gambling marketing and unplanned gambling spend: synthesised findings from two online cross-sectional surveys. *Addict Behav* 2022:107440.
32. Ipsos MORI. *Final Synthesis Report: the effect of gambling marketing and advertising on children, young people and vulnerable adults*. London: GambleAware; 2020.
33. Thomas C, Breeze P, Cummins S, Cornelsen L, Yau A, Brennan A. The health, cost and equity impacts of restrictions on the advertisement of high fat, salt and sugar products across the transport for London network: a health economic modelling study. *Int J Behav Nutr Phys Activ* 2022;**19**:1–12.
34. Yau A, Berger N, Law C, Cornelsen L, Greener R, Adams J, et al. Changes in household food and drink purchases following restrictions on the advertisement of high fat, salt, and sugar products across the Transport for London network: a controlled interrupted time series analysis. *PLoS Med* 2022;**19**:e1003915.
35. Meiksin R, Er V, Thompson C, Adams J, Boyland E, Burgoine T, et al. Restricting the advertising of high fat, salt and sugar foods on the Transport for London estate: process and implementation study. *Soc Sci Med* 2022;**292**:114548.

Public Health 215 (2023) 39–41

Contents lists available at ScienceDirect

Public Health

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

Themed Paper — Commentary

Why time matters when it comes to resilience: how the duration of crisis affects resilience of healthcare and public health leaders

C. Förster [a], N. Füreder [b, *], A. Hertelendy [c, d]

[a] Technical University Chemnitz, Faculty of Economics and Business Administration Germany, Germany
[b] Johannes Kepler University, Sustainable Transformation Management Lab & JKU Business School, Altenberger Strasse 69, 4040 Linz, Austria
[c] Florida International University, Department of Information Systems and Business Analytics, College of Business, Miami, FL, USA
[d] Beth Israel Deaconess Medical Center, Department of Emergency Medicine, Disaster Medicine Fellowship, Boston, MA, USA

ARTICLE INFO

Article history:
Received 8 August 2022
Received in revised form
9 November 2022
Accepted 28 November 2022
Available online 11 January 2023

Keywords:
Healthcare
Public health
Crisis management
Time
Leaders' resilience strategies

ABSTRACT

The meaning of time, especially in crisis, where situations are likely to become even more complex, uncertain, and disruptive, is crucial. Incorporating previous research on organizational crises, organizational resilience, extreme context, and individual resilience, we know that leaders do play a crucial role when it comes to handle adversity in organizations but also that leaders might influence organizational resilience and employee resilience. Intensified by the COVID-19 pandemic, the leaders' ability to effectively deal with a critical situation becomes even more important in healthcare organizations. We argue that time is not only important when it comes to crisis management but also that it is highly significant when it comes to leaders' resilience. Considering the aspect of time implies that different temporal demands, especially regarding the persistence of adversity, require different resilience strategies applied by the leader. Therefore, we call for future research on examining how different leaders' resilience strategies (short term vs. long term) affect crisis management outcomes as well as the resilience in healthcare and public health organizations.

Living in a highly globalized and intertwined VUCA (i.e. volatile, uncertain, complex, and ambiguous) world, the meaning of time, in particular in the context of crisis, where the situation is likely to become even more complex, uncertain, and disruptive,[1] is crucial. Defining organizational crises as "an event perceived by managers and stakeholders as highly salient, unexpected, and potentially disruptive," the meaning of time is characterized both in terms of the unexpectedness of the crisis and "the perceived urgency of the response." Considering the implicit meaning of time in organizational crisis, it is not surprising that time is also incorporated into crisis management and leadership. Derived from the "perceived significance and urgency of crises," "time pressure, risks, and uncertainty" are the conditions underlying leaders' decision-making in crisis. Researchers have conceptualized crisis leadership around the aspect of time by referring to the time before crisis, during crisis, and after crisis,[2] whereby time has been a key element of crisis management in the recent COVID-19 pandemic.[3,4]

Globally, crisis management is crucial in healthcare and public health organizations that are used to operating in risky contexts, where the risks of catastrophes are omnipresent. Although the possibility of crisis is ubiquitous in healthcare, and hospitals, before the COVID-19 pandemic, only single units of these organizations, for example, emergency or intensive care units,[5] needed to deal with routine emergencies and mass casualty incidents. Furthermore, crisis management in these specific units was largely reactive. Along this line, medical teams needed to respond to these events and mobilize resources rapidly, whereby the duration of exposure to such emergencies events was relatively short.[6] However, the emergence of SARS-CoV-2 and the subsequent pandemic changed this by moving hospitals into a temporal organization-wide emergency context, where the "potentiality of crisis" turned into actuality.[5] In this sense, including this shift in context, we call for future research examining the aspect of time regarding different leaders' resilience strategies and how this affects the resilience of healthcare and public health organizations to contribute to the international discourse.[7,8]

Healthcare and public health leaders' resilience

Considering previous research on organizational crises,[9] organizational resilience,[10] extreme context,[6] and individual resilience in the workplace,[11] we know that leaders do play a crucial role

* Corresponding author.
E-mail address: nina.fuereder@jku.at (N. Füreder).

https://doi.org/10.1016/j.puhe.2022.11.024

when it comes to handling adversity in organizations but also influence resilience through a trickle-up (i.e. organizational resilience) and trickle-down effect (i.e. employee resilience).

Intensified by the COVID-19 pandemic, the leaders' ability to effectively deal with a critical situation becomes even more important in healthcare and public health organizations. Derived from Foerster and Duchek's[12] process-oriented model, leaders' resilience can be understood, as three successive stages of the resilience process, namely, precrisis, during crisis, and postcrisis, whereby each stage requires a specific behavior of the leader to master these stages successfully. While, in stage 1, the leader must anticipate critical developments and prepare for potential crises; in stage 2, the leader must cope with an acute crisis. In stage 3, the leader must finally reflect on and learn from the crisis.

Although this dynamic process provides us with some notion on how leaders effectively deal with critical situations, the understanding of time primarily refers to the general division of crisis into the three phases of pre, during, and after crisis[1,2] but leaves us with a lack of understanding about the influence of time within each of these stages. Along this line, longer "coping" phases during crisis, as was the case during the pandemic, might require a different approach by healthcare leaders than during shorter "coping" phases, which are more common in hospitals, especially in terms of non-routine emergencies, such as mass casualty incidents or disasters.

Time and temporality in resilience

As we know that a crisis can last for a long time, temporal conditions for managing crisis become increasingly salient. Ancona et al.[13] define time as "a non-spatial continuum in which events occur in apparently irreversible succession from the past through the present to the future."[14] Time is increasingly considered an important resource in organizational research.[15] Scholars have called for a sharper focus for understanding temporality in organizational relationships, employee interactions, and performance, whereby time has an important influence on organizational and social practices.[16]

Regarding resilience, time has multifactorial impacts relating to anticipating, coping, and how to recover and learn from a crisis. Although healthcare leaders have been threatened by a potential vicious circle of stress endangering the healthcare leaders' resilience long before the pandemic, the COVID-19 pandemic has exacerbated this situation. During the pandemic, healthcare leaders were stressed in multiple ways; they had to handle a high workload, manage many changes in a short time and tolerate a high degree of uncertainty. Therefore, resilience emerged as a prerequisite for overcoming the pandemic not only in a productive but also in a healthy manner.

Incorporating the aspect of time, Bardoel and Drago[17] propose two types of resilience, namely, acceptance resilience and strategic resilience. Originating from the conservation of resources theory, acceptance resilience is built on a "resource-preserving" strategy and is therefore probably more suitable for short-term minor adversity, strategic resilience involves a "resource-enhancing" strategy, and it is thus probably more suitable for major long-term adversity.[17] Although both strategies are successful coping mechanisms, they differ fundamentally in their use of resources. Along this line, individuals who are predominantly used to acceptance resilience will experience declining resources over time, whereas those who rely on strategic resilience will gain resources over time.

Referring to the emergency context in healthcare and public health, where crises are usually short but intensive, it can be assumed that the crisis structure in healthcare might favor a "resource-preserving" acceptance resilience strategy. However, long-term adversity, such as the pandemic, or persisting overwork, where the work demand constantly exceeds the individual resources,[18] which is typical in healthcare and public health organizations, might require "resource-enhancing" strategic resilience strategies. Otherwise, the institutional and organizational constraints are likely to cause the healthcare leaders' resilience to erode throughout time,[19] whereby consequences can be severe. Previous studies suggest that an overload due to psychological job demands can be associated with negative events, such as near misses and injuries.[20] At the individual level, constant stress and overload can be related to both psychological as well as physical health problems, such as heart disease and other chronic ailments.

Conclusion

Building on previous research on resilience, especially on leaders and in the healthcare setting,[8] we argue that time is not only important when it comes to crisis management but is also highly significant when it comes to leaders' resilience, implying that different temporal demands, ely regarding the persistence of adversity, require different resilience strategies applied by the leader. Future research might examine how different leaders' resilience strategies affect crisis management outcomes as well as the resilience in healthcare and public health organizations. As we know that individuals' coping resources decline over time, we must expect that those who stay with acceptance resilience may experience deleterious effects such as burnout or chronic sickness, and those who switch from acceptance to strategic resilience will seek new opportunities for occupational engagement. From a practical point of view, this could end up in precarious leadership shortages where leaders either are on long-term sick leave or, what we already observe, voluntarily relinquishing their leadership positions and returning to clinical practice.

Author statement

Ethical approval

None sought.

Funding

This work emerged from a research project supported by the High Impact Program of the JKU Business School. There are not any ethical, approval, funding and competing interests because of the conceptual nature of this short communication paper.

Competing interests

None declared.

References

1. Bundy J, Pfarrer MD, Short CE, Coombs WT. Crises and crisis management: integration, interpretation, and research development. *J Manag* 2017;**43/6**:1661–92.
2. Wu YL, Shao B, Newman A, Schwarz G. Crisis leadership: a review and future research agenda. *Leader Q* 2021;**32**(6):1015–8.
3. Hertelendy AJ, Chekijian S, McNulty E, Mitchell CL, Grimes JO, Durneva P, et al. Crisis leadership: a case for inclusion in accredited master of public health program curricula. *Publ Health* 2022;**209**:14–8.
4. Kunisch S, Blagoev B, Bartunek JM. Complex times, complex time: the pandemic, time-based theorizing and temporal research in management and organization studies. *J Manag Stud* 2021;**58**(5):1411–5.
5. Hällgren M, Rouleau L, De Rond M. A matter of life or death: how extreme context research matters for management and organization studies. *Acad Manag Ann* 2018;**12**(1):111–53.

6. Hannah ST, Uhl-Bien M, Avolio BJ, Cavarretta FL. A framework for examining leadership in extreme contexts. *Leader Q* 2009;**20**(6):897–919.
7. Yu M, Wen J, Smith SM, Stokes P. Building-up resilience and being effective leaders in the workplace: a systematic review and synthesis model. *Leadership Org Dev J* 2022 [Nr. ahead-of-print].
8. West CP, Dyrbye LN, Sinsky C, Trockel M, Tutty M, Nedelec L, et al. Resilience and burnout among physicians and the general US working population. *JAMA Netw Open* 2020;**3**(7). e209385-e209385.
9. Pearson CM, Clair JA. Reframing crisis management. *Acad Manag Rev* 1998;**23**(1):59–76.
10. Buyl T, Boone C, Wade JB. CEO narcissism, risk-taking, and resilience: an empirical analysis in US commercial banks. *J Manag* 2019;**45**(4):1372–400.
11. Foerster C, Duchek S. Leaders' resilience-A systematic literature review and future research agenda. *Acad Manag Proc* 2018;**1**:13879.
12. Foerster C, Duchek S. Leaders' Resilience: what leaders can learn from the COVID-19 crisis. In: Dhiman S, Marques J, editors. *Leadership after COVID-19 working together toward a sustainable future*. Cham: Springer; 2022. p. 17–30.
13. Ancona DG, Okhuysen GA, Perlow LA. Taking time to integrate temporal research. *Acad Manag Rev* 2001;**26**(4):512–29.
14. Sirén C, Parida V, Frishammar J, Wincent J. Time and time-based organizing of innovation: influence of temporality on entrepreneurial firms' performance. *J Bus Res* 2020;**112**:23–32.
15. Berends H, Antonacopoulou E. Time and organizational learning: a review and agenda for future research. *Int J Manag Rev* 2014;**16**(4):437–53.
16. Langley ANN, Smallman C, Tsoukas H, Van de Ven AH. Process studies of change in organization and management: unveiling temporality, activity, and flow. *Acad Manag J* 2013;**56**(1):1–13.
17. Bardoel EA, Drago R. Acceptance and strategic resilience: an application of conservation of resources theory. *Group Organ Manag* 2021;**46**(4):657–91.
18. Sandhåland H, Oltedal HA, Hystad SW, Eid J. Effects of leadership style and psychological job demands on situation awareness and the willingness to take a risk: a survey of selected offshore vessels. *Saf Sci* 2017;**93**:178–86.
19. Foerster C, Duchek S, Geithner S. Resilienz von Führungskräften in pandemischen zeiten: Stärkung oder Erosion? In: Zerth J, Garkisch M, Nass E, editors. *Leben und Versorgung gestalten nach Corona: lernen und Lehren aus der Krise*. Stuttgart: Kohlhammer; 2022. p. 149–66.
20. Li F, Jiang L, Yao X, Li Y. Job demands, job resources and safety outcomes: the roles of emotional exhaustion and safety compliance. *Accid Anal Prev* 2013;**51**:243–51.