Public Health 202 (2022) 43–48

Contents lists available at ScienceDirect

Public Health

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

Original Research

An off-target scale limits the utility of Short Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS) as a measure of well-being in public health surveys

J. Melin [a, b, *], A. Lundin [c], M. Johansson [a]

[a] RISE Research Institutes of Sweden, Safety and Transport, Measurement Science and Technology, Sweden
[b] The Gothenburg Centre for Person Centred Care (GPCC), University of Gothenburg, Sahlgrenska Academy, Sweden
[c] Global Public Health, Karolinska Institutet, Sweden

ARTICLE INFO

Article history:
Received 3 July 2021
Received in revised form
6 October 2021
Accepted 19 October 2021
Available online 6 December 2021

Keywords:
Well-being
Rasch
Measurement
Targeting
Public health

ABSTRACT

Objectives: To assess the utility and measurement properties for the well-being scale Short Warwick–Edinburgh Mental Well-Being Scale (SWEMWBS) in a Swedish general population survey.
Study design: A cross-sectional survey study.
Methods: Data were retrieved from the 2018 public health survey in Stockholm County, containing a random sample of 22 856 persons stratified to be representative for the municipalities and districts within the region. The data were analyzed according to Rasch Measurement Theory.
Results: Person attribute values are positively skewed (mean 2.32, SD 1.85), with wide gaps in the item threshold attribute values. Overall item fit statistics were acceptable, and person measurement separation reliability was 0.83, indicating three statistically distinct ranges in the estimated well-being values.
Conclusion: While the SWEMWBS items indicated acceptable fit to the Rasch measurement model, targeting of items to sample is skewed toward lower levels of well-being, and there is a ceiling effect. Thus, we suggest a careful reconsideration of SWEMWBS as a tool for use in general public health surveys, especially for assessing change over time and group differences, as there are large measurement uncertainties for the majority of cases when the population as a whole is sampled. We encourage revisions applying a coherent and comprehensive ordinal construct theory for well-being to fill the gaps in the upper end of the SWEMWBS scales' item thresholds. The addition of more challenging items would improve targeting for population-based surveys, increase reliability, and provide more actionable information that could be useful in improving individuals' well-being.

Introduction

Worldwide and across disciplines, there is a growing interest to monitor and follow the development of population mental health and well-being.[1] One of the United Nations sustainability goals (SDG 3.4) states: *By 2030, reduce by one third premature mortality from non-communicable diseases through prevention and treatment and promote mental health and well-being.*[2]

To quote VanderWeele:[3] *In order to make more substantial improvements to the well-being of our world, it is critical that we have the data to measure it.* However, data is not the sole criterion for success in measurement. A scientific approach to measurement requires empirical data based on sound theories, as well as the proper treatment of collected data. For nearly 100 years, since the work of Thurstone in the 1920s, criteria for calibrating sets of items measuring in theoretically defined unit quantities have informed research and practice across dozens of fields in psychology and the social sciences. This includes the work by Danish mathematician Georg Rasch in the 1960s, who based his models on the same underlying principles as physical measurements, developed the Rasch Measurement Theory (RMT) to enable separate measures of person and item attributes scaled on a conjoint interval logit scale.[4] In physical measurement, it is obvious that the properties of the instrument must be independent of the properties of the objects, and vice versa. Similarly, person and item attributes in social and psychological measurements need to be independent of each other, i.e. the specific objectivity proposed by Rasch.[4]

* Corresponding author. Sven Hultins plats 5, 412 58 Gothenburg, Sweden. Tel.: +4610 516 66 39.
E-mail address: Jeanette.melin@ri.se (J. Melin).

https://doi.org/10.1016/j.puhe.2021.10.009

As argued by Salzberger et al.:[5] '*the application of the Rasch model alone does not guarantee better social measurement, if statistical analyses lack a conceptual underpinning*'. The underlying theoretical model is a central foundation to the assessment of a construct, particularly when self-rated responses are our primary source of data. This can be challenging already at the theoretical level. For instance, within the field of well-being, there is an abundance of models that have been developed over the years,[6–8] and the field has yet to find parsimony and comparability. However, a critical point is an acknowledgment that mental constructs are accessible to measurement and in turn relating self-ratings to an abstract linear continuum of 'less to more'.[9] RMT provides a possibility for connecting the empirical evidence with a conceptual – predefined or ad hoc – theory; the item-hierarchy provides '*a compass and interpretable language*',[10] directing the way evidence-based interventions and innovations.

One example of a widely used well-being scale is the *Short Warwick–Edinburgh Mental Well-Being Scale* (SWEMWBS, stemming from the *Warwick–Edinburgh Mental Well-Being Scale* (WEMWBS)[11]) that in previous work has been evaluated to basic measurement criterion according to RMT in public health surveys.[12] The SWEMWBS is used in several national population surveys to estimate well-being and how it develops over time.[13,14] The SWEMWBS is intended to focus on the positive aspects of well-being. Thus, it should be free of ceiling effects in population samples.[11] The SWEMWBS is included in the Swedish biannual public health survey *Hälsa på lika villkor*,[15] but its measurement properties have not previously been evaluated for this population. The purpose of this study was to investigate the utility and measurement properties of the SWEMWBS according to RMT in a Swedish general population sample.

Methods

Participants and data collection

We retrieved data from the 2018 annual public health survey *Hälsa på lika villkor*[15] in Stockholm County, which contains a stratified random sample of 22 856 persons (strata are municipalities and districts within the region). The overall response rate, percentage of eligible respondents, was 39.2%. Design weighted response rate, representing the response if people had the same sampling probabilities were 39.1%. The median age was 54 years (range 16–84), 55% ($n = 12{,}571$) were women and 45% ($n = 10{,}285$) were men, and 81% ($n = 18{,}513$) were born in Sweden and 19% ($n = 4343$) were born in other countries.

Measurement

The SWEMWBS[12] is a 7-item short form of the longer 14-item *Warwick–Edinburgh Mental Well-Being Scale* (WEMWBS).[11] Item generation for the WEMWBS was based on an expert panel considering literature in the field, qualitative focus groups, and psychometric testing by comparison with a related scale, the *Affectometer 2*. The WEMWBS is purported to cover both hedonic and eudemonic (or psychological) aspects of mental health.[11] Item reduction for SWEMWBS was based on RMT analysis to ensure a unidimensional scale;[12] items mainly representing psychological and eudemonic well-being and less related to hedonic well-being or affect.

The respondents are asked to rate each of the seven items in relation to how often they experienced each statement in the past two weeks. The original response alternatives *none of the time, rarely, some of the time, often* and, *all of the time* in the present study were changed to *never, seldom, sometimes, often,* and *always*, which are scored 0 through 4. This version came from the WHO Health Behavior in School-aged Children (HBSC) Sweden.[16] All items use positive wording.

Data analysis

In order to assess the measurement properties of SWEMWBS, the RMT was applied, as '*it is not simply a mathematical or statistical approach, but instead a specifically metrological approach to human based measurement*'.[17] RMT was developed based on the same underlying principles as physical measurements, i.e. separate measures of person and item attributes scaled on a conjoint interval logit scale.[4] In the simplest case with dichotomous responses, it is a logistic regression function:

$$\log\left(\frac{P_{success,i,j}}{1 - P_{success,i,j}}\right) = \theta_i - \delta_j \quad (1)$$

where θ is the person, i is the person attribute value, and δ is the item, and j is the item attribute value. This model enables estimates of the differences between the person and task attributes to meet requirements for invariant comparisons. Overall, one wants to have items that work in the same way for the different target groups. In some cases, certain subgroups may not respond to items in a similar way, i.e. the item/items fail to meet the criteria of invariance. This is tested by Differential Item Functioning (DIF) statistics, as described below.

The dichotomous model has subsequently been expanded to polytomous RMT models by incorporating threshold parameters,[18,19] which we applied to SWEMWBS data, as the questionniare comprises five response categories. The analyses were conducted using the Rasch Unidimensional Measurement Model 2030 (RUMM) software.

The analysis was structured in line with recommendations by Hobart & Cano.[20] Initially, an examination of the relative distributions of the item threshold attribute values to the person attribute values was assessed. In order to understand the implications of the person measures and how those should be used, item threshold-person distributions, Person Separation Index (PSI), and person fit residuals were examined. Specifically:

a) The mean person attribute value indicates whether the sample is off-center relative to the items;
b) The PSI is a reliability indicator where 0 implies all error/noise, 1 implies no error, and the coefficient should be over 0.8 (i.e. corresponding to a person separation index of 2.0, and the definition of three statistically different strata[21]);
c) The person fit residual should lie within −2.5 to +2.5. In addition to those aspects, individual person measurement uncertainties were assessed for the required change in well-being to become detectable. We applied $k = 2$, which approximately corresponds to a 95% confidence interval.

In order to understand the item properties, individually and together, response categories were assessed to see if they worked as intended, and items were studied to see if they exhibited a clinically logical order. Response categories and items were assessed in terms of model fit residuals, chi-square, item characteristic curve (ICC), residual correlation, unidimensionality, and DIF statistics, specifically:

a) Category thresholds should be sequentially ordered, and as the trait level is increasing, the probability of a higher response also increases;[22]
b) The individual item fit residuals should lie between −2.5 and + 2.5; the chi-square values should not be statistically

significant after applying Bonferroni correction; and the dots of the observed class intervals should follow the expected ICC to support a good fit to the model;

c) Construct definition across samples was examined by comparing the item locations obtained in this study with those obtained in two previously published RMT articles analyzing SWEMWBS data;[23,24]
d) The item residual correlations should not exceed 0.20 above the average correlations[25] to support local independence (potential indication of multidimensionality and/or redundancy due to too similarly phrased items);
e) Person attribute values estimated from two subsets of items (derived based on item loadings in the first factor in the principal component analysis of the residuals) were compared by an independent *t*-test where the percentage of persons outside the range of −1.96 to 1.96 should not exceed 5% to support unidimensionality;[26]
f) For DIF, both main effects and interaction effects were assessed for between gender (men; women) and age (≤25; 26–45; 46–65; ≥66 years) and should be non-significant after Bonferroni correction.[27] When significant, DIF was followed by stepwise item splits and repeated analyses.

For avoiding misfit inflated by the large sample size, item chi-squares and DIF were assessed using the whole sample, as well as with adjusted sample sizes for $n = 200$ and $n = 500$.[28]

Results

As shown in Fig. 1 and Table 1, item threshold attribute values are covered by the person attribute values, but not the other way around. Especially, person attribute values are positively skewed, and there are gaps (indicated by the blue arrows in Fig. 1) in the item threshold attribute values. The PSI was 0.83, indicating three statistically distinct ranges in the estimated well-being values.

In total, 3011 (14.2%) persons had fit residuals ≤ -2.5, i.e. having low variation in their response pattern, and 103 (0.5%) persons had fit residuals ≥2.5, i.e. having irrational response pattern. In total, 1574 (6.9%) persons responded *always* (i.e. highest response option giving extremes) on all seven items.

None of the items had disordered thresholds, indicating that the response options worked as intended, i.e. progressively more often used as overall wellbeing increased (Fig. 2). Table 2 reports item mean attribute values, associated measurement uncertainties, fit residuals, and ChiSq. All except items 2 (*I've been feeling useful*) and 3 (*I've been feeling relaxed*) had fit residuals outside the desired range ±2.5 and significant ChiSq. However, by assessing adjusted sample sizes and visual inspection of the dots of the class intervals in relation to the ICC, it was evident that the misfit (fit residuals and ChiSq) was inflated by the large sample size.

Table 1
Summary of measurement properties for SWEMWBS. Total person $n = 22\,856$, extremes $n = 1574$.

Measurement property	Values
Person attribute values mean (SD)	2.32 (1.85)
Person attribute mean 2SE	1.47
Person attribute values range with extremes	−5.17 to 6.20
Person attribute values range without extremes	−4.33 to 5.28
Person attribute PSI with extremes	0.83
Person attribute PSI without extremes	0.85
Item attribute values range	−0.57 to 0.53
Item thresholds attribute values range	−3.37 to 4.51

PSI, Person Separation Index; SWEMWBS, Short Warwick–Edinburgh Mental Well-Being Scale.

As shown in Table 3, mean item attributes hierarchy slightly deviated from the two previous Rasch studies on SWEMWBS; item 2 (*I've been feeling useful*) had a relatively higher item attribute value in the present Swedish cohort compared to a UK cohort[23] and an Australian cohort.[24]

The relative cut-off, i.e. 0.2 above the average correlation of −0.16, for item residuals was 0.04. A closer inspection of the residual matrix showed that one item pair correlated above the cut-off; item 4 (*I've been dealing with problems well*) and item 5 (*I've been thinking clearly*) residuals correlated to 0.13. Thus, items were overall locally independent from each other.

Person attribute values derived from the positively and negatively loading items in the first factor in the principal component analysis of the fit residuals were compared by an independent *t*-test; 7% of the person attribute values were outside the desirable ±1.96, i.e. slightly above the criterion of 5% but within the confidence interval, thus supporting unidimensionality.

Uniform DIF was present between men and women, as well as between age groups. It was significant for all comparisons on all

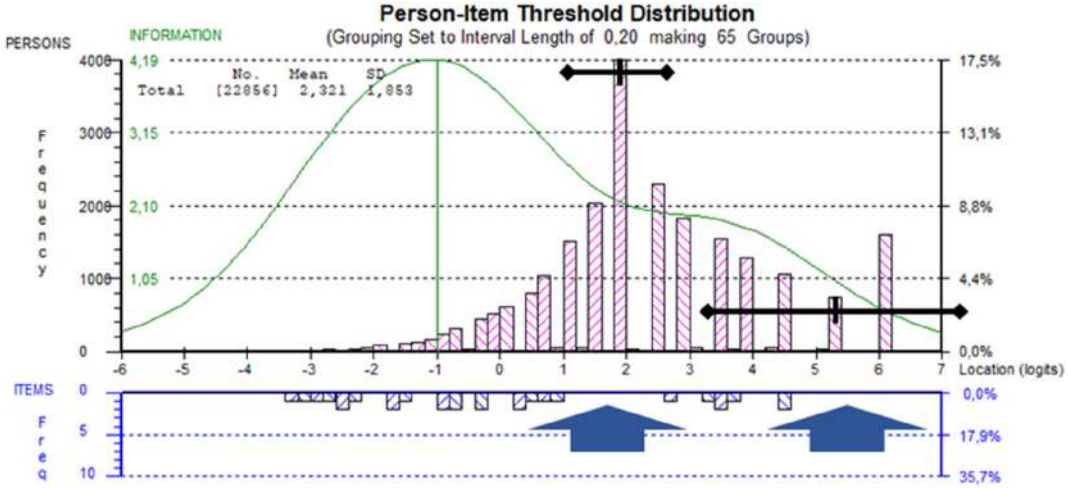


Fig. 1. Person-item threshold histograms SWEBSWBS. Pink upper bars show person attribute values distributions, and blue lower bars item threshold attribute values distributions scaled on a conjoint logit scale. Low values indicate lower well-being, and high values indicate higher well-being. The green curve show where most information about the persons is provided and are inverse functions of the measurement standard errors (SE). Horizontal error bars correspond to 2 SE. Blue arrows indicate gaps in the item threshold attribute values compared to the person attribute values. SWEMWBS, Short Warwick-Edinburgh Mental Well-Being Scale. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.).

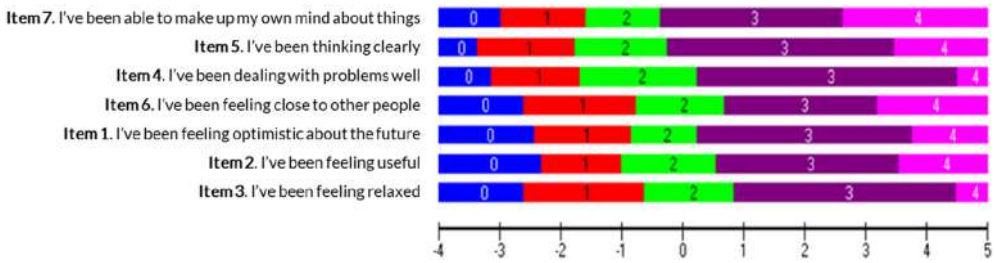


Fig. 2. Item threshold map for SWEMWBS where 0 = *never*, 1 = *seldom*, 2 *sometimes*, 3 = *often*, and 4 = *always*. SWEMWBS, Short Warwick-Edinburgh Mental Well-Being Scale.

Table 2
Item fit statistics for SWEMWBS.

Item		Location	2SE	Fit Residuals	ChiSq	$n = 21{,}158$ Prob	$n = 200$ Prob	$n = 500$ Prob
7	I've been able to make up my own mind about things	−0.57	0.02	**4.95**	270.20	**0.00**	0.98	0.70
5	I've been thinking clearly	−0.47	0.03	**−17.37**	274.10	**0.00**	0.98	0.69
4	I've been dealing with problems well	−0.02	0.03	**−17.47**	400.49	**0.00**	0.92	0.40
6	I've been feeling close to other people	0.14	0.02	**3.96**	109.38	**0.00**	1.00	0.98
1	I've been feeling optimistic about the future	0.19	0.02	**−7.94**	59.80	**0.00**	1.00	1.00
2	I've been feeling useful	0.21	0.02	−0.59	30.14	**0.00**	1.00	1.00
3	I've been feeling relaxed	0.53	0.02	−1.78	187.52	**0.00**	0.99	0.88

Bolded numbers indicate misfit; Fit Residuals ±2.5 or significant ChiSq. SWEMWBS, Short Warwick–Edinburgh Mental Well-Being Scale.

items except item 4 (*I've been dealing with problems well*) among different age groups. However, the number of significant DIF were reduced when assessing adjusted sample sizes ($n = 200$ no significant DIF, $n = 500$ item 3 for gender and items 2 and 3 for age), and item splits (i.e. item divided for person factors) could have resolved the DIF without significantly impacting person attribute value. Person attribute values with and without item splits showed correlation coefficients above 0.99, supporting little overall effect of these items being perceived differently between the groups.

Discussion

This study has demonstrated an acceptable fit of SWEMWBS to the basic Rasch measurement model. However, the findings raise several concerns about the utility of SWEMWBS as a measure of well-being in public health surveys. As shown in Fig. 1, the best measurement region is indicated by the green information curve that overlaps with the low well-being end of the spectrum. Moreover, the highly positive skewed person attribute values imply large measurement uncertainties for most of the population. As such, this requires large changes in well-being to be detectable and not biased due to measurement noise. Moreover, there were many extremes on the positive side, with 1574 (6.9%) persons responding *always* (i.e. highest response option) on all seven items, and it is obvious that we know very little about their well-being, and it would be impossible for them to show improvement on the scale. The SWEMWBS has previously been contended as being free of ceiling effects in general population samples,[11] but our analysis does not support this claim. We suggest careful consideration in the use and interpretation of SWEMWBS in public health surveys, especially for assessing change, due to the large mistargeting between sample and items and the ceiling effect observed. Although, it might be used to screen for low well-being or with groups with lower levels of well-being.

Early work on SWEMWBS took a step in the right direction to a scientific approach for measurement when addressing the measurement properties according to RMT[12] rather than classical test theory.[11] However, further development to a scientific measurement approach has to a great extent been absent for the SWEMWBS, as well as well-being measures in general. There are current recommendations for measuring well-being in public surveys claimed to be based on the state of knowledge,[29] that lack a scientific measurement approach. The recommendations include having four or fewer generic questions. Suggestions of a few items are probably related to the lack of insight into the issues with measurement uncertainties since ordinal responses are typically statistically treated as having interval scale properties. However, measurement uncertainties reflect the lack of exact knowledge of the value of the measurand (e.g. person attribute values and item attribute values) and characterizes the dispersion of the values that

Table 3
Item δ attributes values for the Swedish cohort compared with a UK cohort[16] and an Australian cohort.[17] The UK study included the same seven items while the Australian study included the same seven items, as well as three additional items from WEBWBS.

Item		Region Stockholm 2018		Bartram et al. 2013		Hougthon et al. 2017	
		Location	2SE	Location	2SE	Location	2SE
7	I've been able to make up my own mind about things	−0.57	0.02	−1.07	0.08	−0.85	0.07
5	I've been thinking clearly	−0.47	0.03	−0.66	0.08	−0.48	0.07
4	I've been dealing with problems well	−0.02	0.03	0.08	0.08	−0.19	0.07
6	I've been feeling close to other people	0.14	0.02	0.11	0.06	−0.09	0.07
1	I've been feeling optimistic about the future	0.19	0.02	0.48	0.06	0.47	0.06
2	I've been feeling useful	0.21	0.02	**0.00**	**0.08**	**−0.14**	**0.06**
3	I've been feeling relaxed	0.53	0.02	1.06	0.06	0.69	0.06

Bolded numers indicate a realtively higher item attribute compared to the Swedish cohort. WEMWBS, Warwick–Edinburgh Mental Well-Being Scale.

reasonably could be attributed to it.[30] Few items and/or mistargeted items imply low measurement precision.[31] More items mean that we get more information about each respondent and especially if the number of well-targeted items increases the measurement uncertainties for estimates of person attribute values are reduced. Although the measurement uncertainties will always be greater at the upper (and lower) end of a scale as the model anticipates infinity in both directions,[21] and limits reliable decisions on general well-being in public health surveys for people with moderate to high well-being.

In order to accommodate the mistargeting and to further advance possibilities for an optimal well-being scale to be used in public health surveys, we would argue for addressing the issue of a coherent ordinal construct theory[32] for well-being. First, drawing from Fig. 1, the person-item thresholds maps the well-being terrain by showing the person's overall well-being relative to the item properties and should be used to extend the understanding of what characterizes low, medium, and high levels of well-being. Second, as there is a gap in item thresholds, as well as lacking items/item thresholds at the upper end of the scale, we suggest that the SWEMWBS is extended with more challenging items to improve its targeting for population-based surveys. Adding items with higher levels of difficulty should ideally be driven by the ordinal construct theory and the target population's understanding of what characterizes higher levels of well-being. However, the benefit of adding items needs to be balanced with the possible risk of reduced response rates. This is particularly true for groups with low levels of awareness and engagement (see further below). Minor studies where new sets of items are tested, based a coherent ordinal construct theory, to demonstrate that the 'right items' are added is recommended before conducting large public health surveys.

In addition to our findings, the previous studies on WEMWBS and SWEMWBS could also be used to extend the ordinal construct theory. When comparing the results in this study with those, there are at last two critical issues that need attention and further cross-country evaluation. First, as shown in Table 3, item 2 (*I've been feeling useful*) had a relatively higher item attribute value in the present Swedish cohort compared to the UK and Australian cohorts.[23,24] This raises questions whether feeling useful has a different meaning in different cultures and/or whether there might be translation issues. Second, the Swedish cohort has more extreme responses and shows greater mistargeting between person and items compared to the other studies. This could be explained by different sample characteristics, cultural differences, or a different response scale.

Apart from the suggestion of adding (or replacing) items, this also needs to be considered in the context of the fact that the SWEMWBS might represent a few items out of hundreds of items included in public health surveys. A concern apparent in our study of SWEMWBS, but less discussed in the literature, is the large number of persons ($n = 3011$, 14.2%) having low variation in their response pattern. This raises questions such as; how valid are person responses when they tick the same response option for all seven items? Did the person even read the item, or is this a matter of acquiescence, i.e. agreement regardless of item content? Or is this a consequence of the response burden in lengthy public health surveys?

Since this study only used cross-sectional data, it was not possible to make evaluations of item stability over time or sensitivity to change. A study by Maheswaran and colleagues[33] claimed that the '*WEMWBS is responsive to changes occurring in a wide range of mental health interventions undertaken in different populations*'. That study did not use person attributes derived from Rasch-transformed interval level measures, which makes the claim difficult to assess. Furthermore, we would encourage proper assessments of measurement uncertainties derived from an uncertainty budget[34] and not only standard deviations for group-level analyses. This is particularly important, as shown in this study large individual measurement uncertainties in the upper end of SWEMWBS.

A major strength of this study is that the data used stem from a public health survey with respondents stratified to be representative of the municipalities and districts within the Stockholm region. However, it is not representative of the whole Swedish population. Moreover, at this stage, it was not possible to control for other person characteristics such as health, family situation, living conditions, employment, or income, but it is required in further research.

Conclusions

Despite an acceptable fit of SWEMWBS to the basic Rasch model, due to the large mistargeting between sample and items and the ceiling effects, there are limitations with of the utility of SWEMWBS as a measure of well-being in public health surveys. Thus, we suggest careful consideration if SWEMWBS is used in general public health surveys, especially for assessing change over time and group differences, as there are large individual measurement uncertainties for the majority of respondents. We encourage addressing a coherent ordinal construct theory for well-being, as well as filling the gaps and upper end of the SWEMWBS scales' item thresholds with more challenging items to improve its targeting for population-based surveys.

Author statements

Acknowledgments

The authors would like to give a particular thanks to Fredrik Lindencrona Region Stockholm/Swedish Association of Local Authorities and Regions for encouraging this work and prof. William P. Fisher, Jr., Living Capital Metrics LLC, Sausalito CA USA/BEAR Center, Grad. School Education, UC Berkeley/RISE, Research Institutes of Sweden, Gothenburg, Sweden for valuable and supportive discussion about measurement sciences.

Ethical approval

Ethical approval was obtained from the Swedish Ethical Review Authority, Dnr 2021-01646.

Funding

Part of this work was supported by the innovation partnership between Region Stockholm and RISE and part by the RISE platform *Centre for Categorically Based Measures.*

Competing interests

None declared.

References

1. OECD Better Life Initiative. *Measuring well-being and progress*. Paris: OECD; 2020 [Internet] [cited 2021 Jan 15]. (How's Life?). Available from: http://www.oecd.org/statistics/better-life-initiative.htm.
2. United Nations. *The 17 goals. Department of economic and social affairs - sustainable development* [Internet] [cited 2020 Dec 21]. Available from: https://sdgs.un.org/goals.
3. VanderWeele TJ. *The global flourishing study*. Psychology Today; 2020 [Internet] [cited 2020 Dec 21]. Available from: https://www.psychologytoday.com/intl/blog/human-flourishing/202005/the-global-flourishing-study.

4. Rasch G. *Probabilistic models for some intelligence and attainment tests*. Danmarks Paedagogiske Institut; 1960. p. 184.
5. Salzberger T, Cano S, Abeth-Webb L, Afolalu E, Chrea C, Weitkunat R, et al. Addressing traceability of self-reported dependence measurement through the use of crosswalks. *Measurement* 2021 May:109593.
6. Huppert FA. The state of wellbeing science: concepts, measures, interventions, and policies. In: Cooper CL, editor. *Wellbeing*. Chichester, UK: John Wiley & Sons, Ltd; 2014. p. 1–49 [Internet] [cited 2020 Jun 1]. Available from: http://doi.wiley.com/10.1002/9781118539415.wbwell036.
7. Keyes CLM. Happiness, flourishing, and life satisfaction. In: *The wiley blackwell encyclopedia of health, illness, behavior, and society*. American Cancer Society; 2014. p. 747–51 [Internet] [cited 2020 Mar 9]. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118410868.wbehibs454.
8. Ryff CD. Psychological well-being revisited: advances in the science and practice of Eudaimonia. *Psychother Psychosom* 2014;**83**(1):10–28.
9. Tesio L. Measuring behaviours and perceptions: rasch analysis as a tool for rehabilitation research. *J Rehabil Med* 2003 Jan 1;**35**(3):105–15.
10. Barbic SP, Cano SJ, Mathias S. The problem of patient-centred outcome measurement in psychiatry: why metrology hasn't mattered and why it should. *J Phys Conf Ser* 2018 Jun;**1044**:012069.
11. Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, et al. The Warwick-Edinburgh mental well-being scale (WEMWBS): development and UK validation. *Health Qual Life Outcome* 2007 Nov 27;**5**(1):63.
12. Stewart-Brown S, Tennant A, Tennant R, Platt S, Parkinson J, Weich S. Internal construct validity of the Warwick-Edinburgh mental well-being scale (WEMWBS): a Rasch analysis using data from the scottish health education population survey. *Health Qual Life Outcome* 2009 Dec;**7**(1):15.
13. Danish Regions. *The Danish national health survey*. 2017 [Internet] [cited 2021 Jan 15]. Available from: https://www.danishhealthdata.com/find-health-data/Danskernes-sundhed.
14. NHS Digital. *Health survey for England*. NHS digital; 2019 [Internet] [cited 2021 Jan 15]. Available from: https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england.
15. Folkhälsomyndigheten. *Nationella folkhälsoenkäten – Hälsa på lika villkor*. 2020 [Internet] [cited 2021 Jan 15]. Available from: http://www.folkhalsomyndigheten.se/folkhalsorapportering-statistik/om-vara-datainsamlingar/nationella-folkhalsoenkaten/.
16. HBSC. *Health behaviour in school-aged children - world health organization collaborative cross-national survey*. 2021 [Internet] [cited 2021 Oct 5]. Available from: http://www.hbsc.org/.
17. Pendrill L. Man as a measurement instrument. *NCSLI Meas* 2014 Dec;**9**(4):24–35.
18. Andrich D. A rating formulation for ordered response categories. *Psychometrika* 1978 Dec 1;**43**(4):561–73.
19. Wright BD, Masters GN. In: *Rating scale analysis*. Chicago: Mesa Press; 1982. p. 206.
20. Hobart JC, Cano SJ. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. *Health technol assess* 2009 Feb;**13**(12) [Internet] [cited 2020 Jan 13]. Available from: https://www.journalslibrary.nihr.ac.uk/hta/hta13120/.
21. *Standard errors and reliabilities: Rasch and raw score* [Internet] [cited 2020 Dec 21]. Available from: https://www.rasch.org/rmt/rmt204f.htm.
22. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care Res* 2007 Dec 15;**57**(8):1358–62.
23. Bartram DJ, Sinclair JM, Baldwin DS. Further validation of the Warwick-Edinburgh mental well-being scale (WEMWBS) in the UK veterinary profession: Rasch analysis. *Qual Life Res* 2013 Mar;**22**(2):379–91.
24. Houghton S, Wood L, Marais I, Rosenberg M, Ferguson R, Pettigrew S. Positive mental well-being: a validation of a rasch-derived version of the Warwick-Edinburgh mental well-being scale. *Assessment* 2017 Apr;**24**(3):371–86.
25. Marais I, Andrich D. Formalizing dimension and response violations of local independence in the unidimensional Rasch model. *J Appl Meas* 2008;**9**(3):200–15. PMID: 18753691.
26. Smith RM. A comparison of methods for determining dimensionality in Rasch measurement. *Struct Equ Model Multidiscip J* 1996 Jan 1;**3**(1):25–40.
27. Hagquist C. Explaining differential item functioning focusing on the crucial role of external information – an example from the measurement of adolescent mental health. *BMC Med Res Methodol* 2019 Sep 5;**19**(1):185.
28. Hagell P, Westergren A. Sample size and statistical conclusions from tests of fit to the Rasch model according to the Rasch unidimensional measurement model (RUMM) program in health outcome measurement. *J Appl Meas* 2016;**17**(4):416–31.
29. VanderWeele TJ, Trudel-Fitzgerald C, Allin P, Farrelly C, Fletcher G, Frederick DE, et al. Current recommendations on the selection of measures for well-being. *Prev Med* 2020 Apr;**133**:106004.
30. *JCGM_200_2012.pdf* [Internet] [cited 2020 Dec 18]. Available from: https://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf.
31. Sarstedt M, Diamantopoulos A, Salzberger T. Should we use single items? Better not. *J Bus Res* 2016 Aug;**69**(8):3199–203.
32. Wilson M. In: *Constructing measures: an item response modeling approach*. Mahwah, N.J: Lawrence Erlbaum Associates; 2005. p. 228.
33. Maheswaran H, Weich S, Powell J, Stewart-Brown S. Evaluating the responsiveness of the Warwick Edinburgh mental well-being scale (WEMWBS): group and individual level analysis. *Health Qual Life Outcome* 2012 Dec 27;**10**(1):156.
34. Fisher WP, Stenner AJ. Towards an alignment of engineering and psychometric approaches to uncertainty in measurement: consequences for the future. In: Corletto C, editor. *18th international congress of metrology*. Paris, France: EDP Sciences; 2017. p. 12004 [Internet] [cited 2020 Dec 2]. Available from: http://cfmetrologie.edpsciences.org/10.1051/metrology/201712004.

Public Health 202 (2022) 121–130

Contents lists available at ScienceDirect

Public Health

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

Original Research

Breast cancer incidence by age at discovery of mammographic abnormality in women participating in French organized screening campaigns

C. Balamou [a, *, f], A. Koïvogui [b, f], R. Rymzhanova [c], S. Cornelis [d], C. Rodrigue-Moulinie [a], N. Sellier [e]

[a] Office De Lutte Contre Les Cancers (ODLC-01), CRCDC-AURA, Site de L'Ain, 12 Rue de La Grenouillère, 01000 Bourg-en-Bresse, France
[b] Comité Départemental des Cancers (CDC-93), CRCDC-IDF, Site de Seine-Saint-Denis, 41 Avenue de Verdun, 93146 Bondy, France
[c] Association pour le Dépistage des Cancers (ADECA-FC), CRCDC-Bourgogne-Franche-Comté, Site de Franche-Comté, 3 Rue Paul Bert, 25000 Besançon, France
[d] Association Icaunaise de Dépistage Du Cancer (AIDEC), CRCDC-Bourgogne-Franche-Comté, Site de L'Yonne, 3 Rue Général de Billy, Saint-Georges-sur-Baulche, Yonne 89000, France
[e] APHP, Service Radiologie, Hôpital Jean-Verdier, Hôpitaux Universitaires de Paris-Seine-Saint-Denis, 93146 Bondy, France

ARTICLE INFO

Article history:
Received 6 July 2021
Received in revised form
4 November 2021
Accepted 14 November 2021
Available online 21 December 2021

Keywords:
Breast cancer screening
Mammography
Mammographic abnormality
Breast cancer incidence

ABSTRACT

Objective: Statistical modeling was already predicted the occurrence/prognosis of breast cancer from previous radiological findings. This study predicts the breast cancer risk by the age at discovery of mammographic abnormality in the French breast cancer screening program.
Study design: This was a cohort study.
Methods: The study included 261,083 women who meet the inclusion criteria: aged 50–74 years, living in French departments (Ain, Doubs, Haute-Saône, Jura, Territoire-de-Belfort, and Yonne), with at least two mammograms between January 1999 and December 2017, of which the first was 'normal/benign'. The incidence of each abnormality (microcalcifications, spiculated mass, obscured mass, architectural distortion, and asymmetric density) was first estimated, then the breast cancer risk was predicted secondly according to the age at discovery of each mammographic abnormality, using an actuarial life table and a Cox model.
Results: Overall breast cancer (6326 cases) incidence was 3.3 (3.0; 3.1)/1000 person-years. The breast cancer incidence increased proportionally with the discovery age of the speculated mass and microcalcifications. The incidence was twice as high when the spiculated mass age of discovery was ≥70 (12.2 [10.4; 14.4]) compared with age 50–54 years (5.8 [5.1; 6.7]). Depending on the spiculated mass discovery age, the breast cancer risk increased by at least 40% between the age groups 55–59 years (1.4 [1.0; 1.8]) and ≥70 years (2.4 [1.9; 3.3]). Whatever the abnormality, the incidence of breast cancer was higher when it was present in only one breast.
Conclusion: The study highlights a stable incidence of breast cancer between successive mammograms, an increased risk of breast cancer with the finding age of spiculated mass and microcalcifications. The reduced delay between the abnormality discovery date and the breast cancer diagnosis date would justify a specific follow-up protocol after the finding of these two abnormalities.

Introduction

Breast cancer (B-cancer) is the world's most common and severe cancer in women: it would be responsible for 11.6% of cancers and 6.6% of deaths from cancer in women.[1] A national coordinated B-cancer screening program (BCSP) has existed in France since 2004.[2,3] The BCSP targets women aged 50–74 years without any

* Corresponding author. Tel.: +33 04 74 45 30 31.
E-mail addresses: christian.balamou@aura.mssante.fr, cbalamou@yahoo.fr (C. Balamou), aakoivogui@live.fr (A. Koïvogui).
[f] First authors with equal contributions.

https://doi.org/10.1016/j.puhe.2021.11.012
0033-3506/

other risk factor except their age and offer them a clinical breast examination and a mammography screening by a licensed radiologist once every 2 years.[2,4] All mammograms considered as normal/benign in a first reading session are submitted to a second reading session.[3]

The BCSP's certified radiologists use the Breast Imaging Reporting System and Data System classification.[5] The radiological findings that support this classification are indicative of certain locations of B-cancer.[6–11] Microcalcifications are present in about 30% of all malignant breast lesions.[10] The positive predictive value of malignancy varies according to the mass contour.[6] Although there are several other risk factors,[7,12,13] it is established that B-cancer can occur without mammographic finding, especially in the case of radiologically dense breasts.[14]

Statistical modeling was already predicted the occurrence of B-cancer or its prognosis from previous radiological findings. A few studies converged toward a positive association between the risk score calculated on previous mammograms and image-detected B-cancer at screening mammograms.[15–18] In addition to the reduced sample size, these studies did not describe the B-cancer risk according to the age at discovery of a radiological abnormality. Likewise, the risk of occurrence of contralateral cancer according to the primary cancer has been documented,[16–20] but the impact of the sequence of appearance of mammographic abnormalities on this risk remains poorly documented to our knowledge.

To perfect the interpretation of these mammographic abnormalities in the daily practice of radiologists, the contribution of artificial intelligence and deep learning is increasingly emphasized.[21–25] Certainly, if there is one field of medicine in which artificial intelligence will offer many advances, it is of course the fields of prevention and screening.[26] However, despite the significant contribution of these new technologies, the conventional reading of mammograms by radiologists is still essential.[21,27] In addition to the discussion on their ethical and legal aspects,[28] there is no algorithm that has beyond doubt been proven to outperform double reporting by two certified breast radiologists.[21] To develop other more efficient algorithms, large databases are needed.[23,24]

Pending the availability of such algorithms, the present study intends to alert radiologists by predicting B-cancer risk by the age at discovery of mammographic abnormalities.

Methods

Study context

The study consisted of a follow-up of 261,083 women, aged 50–74 years, living in six French Departments (Ain, Doubs, Haute-Saône, Jura, Territoire-de-Belfort, and Yonne). These women had at least two screening mammograms between 1 January 1991 and 31 December 2017, of which the first was considered as 'normal/benign'. In the study departments included, screening campaigns began before 2004, full coverage date of the BCSP.[2]

The study excluded (1) women with less than two participations in the BCSP (90,274/351,357) and (ii) women with diagnosis of B-cancer at their first participation (3417/9743 women who had at least one occurrence of B-cancer during the study period).

The definitions of the mammographic abnormalities studied (microcalcifications, spiculated mass, obscured mass, architectural distortion, and asymmetric density) are standardized in the BCSP. However, the study kept the description of microcalcification only if an immediate (or deferred) diagnostic workup confirmed the presence of microcalcification. The incidence of each mammographic abnormality has been described; subsequently, the B-cancer incidence has been described according to the age at discovery of each mammographic abnormality.

BCSP organization

Women with a high risk of B-cancer[3,5] were not eligible in the BCSP because they benefit from specific and annual monitoring.

In the six departments with a target population of 167,401 women in 2017, BCSP management structures (BCSP structure) were in charge, the organization of screening campaigns following BCSP specifications.[3,5] Before each campaign, a list of women was proved and updated according to information from health insurance plans. These women were invited by regular mail at their 50th birthday (first invitation) and then every 2 years after a negative mammogram (subsequent invitation) until the age of 74 years. The letter of invitation allows each woman to have a mammogram (two frontal and two oblique external incidences) in one of the BCSP's certified radiological centers.

In each radiological center, the first reading session of a mammogram was performed on a hard copy or a screen display. The radiologist first reader also had to collect sociodemographic, clinical, and radiological information on a standardized form. Using the printed films, the second reading session, or even a third consensual or expert reading session, was carried out in the BCSP structure.

Campaign dynamics

In the analysis of each mammographic abnormality incidence, the start date of follow-up was the date of the first screening mammogram. This follow-up was censored on the abnormality discovery date or on the study end date if the absence of abnormality in each breast.

In the analysis of the B-cancer incidence, the start date of follow-up was the date of the first screening mammogram. The follow-up was censored on the end date (absence of B-cancer) or on the B-cancer occurrence date (i.e. date of the mammogram that initiated a B-cancer diagnostic procedure) or on the date of interval B-cancer diagnosis in the examined breast. The follow-up was censored on the end date in all other cases, for example, loss to follow-up, relocation, refusal to participate, age >74 years, or death from causes other than B-cancer.

The study distinguished the follow-up of the left breast from that of the right breast. The end date of follow-up was the date of the last mammogram. Women aged <73 years were considered 'lost to follow-up' whenever the date of the last mammogram was before 2015 because they could have had another mammogram before age 74 years. The end of follow-up criterion was the B-cancer diagnosis.

Data collection and factors studied

The data analyzed were extracted from the databases of the BCSP structure. These databases were daily enriched by BCSP partners (health insurance plan, radiologists, pathologists, oncologists, surgeons, gynecologists, and general practitioners).

Regarding B-cancer diagnosis, the study adopted the C50 code (10th version) of the WHO International Classification of Diseases (ICD-10):[29] (1) ductal carcinoma in situ, lobular carcinoma in situ, and nipple Paget's disease were classified as 'adenocarcinoma in situ' (TIS); (2) infiltrating/invasive ductal or lobular carcinomas were classified as 'infiltrating adenocarcinoma' (ADK-I); and (3) all other malignant tumors (papillary, tubular, mucinous, medullary, etc.) were classified as 'rare form'. Was considered first reading cancer (R1-Cancer) when the diagnostic process was started after a positive mammographic in the first reading session (ACR in 0,3,4,5).

Table 1
Cumulative incidence of radiographic abnormalities according to the characteristics of women at baseline.

Characteristics at baseline	Microcalcification			Spiculated mass			Obscured mass			Asymmetric density			Architectural distortion		
	Case (Tp)	C−In [CI 95%]	*P*	Case (Tp)	C−In [CI 95%]	*P*	Case (Tp)	C−In [CI 95%]	*P*	Case (Tp)	C−In [CI 95%]	*P*	Case (Tp)	C−In [CI 95%]	*P*
Left breast															
Overall (n = 260,825)[a]	15,257 (7.1)	8.2 [8.1; 8.3]		12,016 (7.2)	6.4 [6.3; 6.5]		65,505 (6.1)	41.1 [40.8; 41.4]		6365 (7.3)	3.3 [3.2; 3.4]		9681 (7.2)	5.2 [5.1; 5.3]	
Age (year) 1st mammogram			*			*			*			*			*
50−54 (n = 125,261)	7037 (7.0)	8.1 [7.9; 8.3]		5659 (7.1)	6.4 [6.2; 6.6]		29,424 (6.1)	38.8 [38.3; 39.2]		3423 (7.2)	3.8 [3.7; 4.0]		4322 (7.0)	4.9 [4.8; 5.1]	
55−59 (n = 51,186)	3482 (8.6)	7.8 [7.5; 8.0]		2882 (8.9)	6.3 [6.1; 6.6]		14,914 (7.3)	39.7 [39.1; 40.3]		1367 (9.1)	2.9 [2.8; 3.1]		2286 (8.9)	5.0 [4.8; 5.3]	
60−64 (n = 40,727)	2645 (8.2)	8.0 [7.7; 8.3]		2129 (8.3)	6.3 [6.1; 6.6]		11,650 (6.8)	41.9 [41.1; 42.7]		1032 (8.4)	3.0 [2.8; 3.2]		1674 (8.2)	5.0 [4.7; 5.2]	
65−69 (n = 29,311)	1556 (5.6)	9.5 [9.0; 9.9]		992 (5.7)	5.9 [5.8; 6.3]		6858 (4.9)	47.8 [46.7; 49.0]		450 (5.8)	2.7 [2.4; 2.9]		1058 (5.7)	6.4 [6.0; 6.8]	
≥70 (n = 14,340)	537 (2.9)	13.0 [12.0; 14.2]		354 (2.9)	8.5 [7.7; 9.4]		2659 (2.6)	71.4 [68.8; 74.2]		93 (2.9)	2.2 [1.8; 2.7]		341 (2.9)	8.2 [7.4; 9.1]	
Breast density 1st mammogram			*			*			*			*			*
Type I (n = 35,040)	888 (7.2)	3.5 [3.3; 3.7]		1147 (7.2)	4.5 [4.3; 4.8]		8764 (6.1)	41.1 [40.3; 42.0]		582 (7.3)	2.3 [2.1; 2.5]		942 (7.2)	3.7 [3.5; 4.0]	
Type II (n = 161,222)	8479 (7.1)	7.4 [7.3; 7.6]		7412 (7.1)	6.4 [6.3; 6.6]		41,122 (6.0)	42.5 [42.1; 42.9]		4057 (7.2)	3.5 [3.4; 3.6]		6096 (7.1)	5.3 [5.2; 5.5]	
Type III (n = 58,524)	5192 (7.2)	12.3 [12.0; 12.7]		3156 (7.5)	7.2 [7.0; 7.5]		14,377 (6.4)	38.7 [38.0; 39.3]		1613 (7.6)	3.6 [3.5; 3.8]		2463 (7.4)	5.7 [5.4; 5.9]	
Type IV (n = 6039)	698 (7.2)	16.1 [14.9; 17.3]		301 (7.7)	6.5 [5.8; 7.3]		1242 (6.8)	30.3 [28.6; 32.0]		113 (7.8)	2.4 [2.0; 2.9]		180 (7.7)	3.9 [3.3; 4.5]	
HRT 1st mammogram			*			*			*			*			*
No/U (n = 227,089)	12,711 (6.9)	8.1 [7.9; 8.2]		9977 (7.0)	6.2 [6.1; 6.4]		55,875 (6.0)	41.2 [40.9; 41.6]		5503 (7.1)	3.4 [3.3; 3.5]		8100 (7.0)	5.1 [5.0; 5.2]	
Yes (n = 33,740)	2546 (8.3)	9.0 [8.7; 9.4]		2039 (8.5)	7.1 [6.8; 7.4]		9630 (7.1)	40.4 [39.6; 41.3]		862 (8.7)	2.9 [2.7; 3.1]		1581 (8.5)	5.5 [5.3; 5.8]	
Right breast															
Overall (n = *260,854*)[b]	15,040 (7.1)	8.1 [8.0; 8.2]		11,449 (7.2)	6.1 [6.0; 6.2]		62,932 (6.2)	39.2 [38.9; 39.5]		6309 (7.3)	3.3 [3.2; 3.4]		9332 (7.2)	5.0 [4.9; 5.1]	
Age (year) 1st mammogram			*			*			*			*			*
50−54 (n = 125,254)	6878 (7.0)	7.9 [7.7; 8.1]		5454 (7.1)	6.2 [6.0; 6.3]		28,339 (6.1)	37.1 [36.6; 37.5]		3391 (7.2)	3.8 [3.7; 3.9]		4204 (7.0)	4.8 [4.6; 4.9]	
55−59 (n = 51,186)	3460 (8.8)	7.7 [7.5; 8.0]		2667 (8.9)	5.8 [5.6; 6.1]		14,128 (7.4)	37.2 [36.6; 37.8]		1422 (9.1)	3.1 [2.9; 3.2]		2173 (8.9)	4.8 [4.6; 5.0]	
60−64 (n = 40,748)	2648 (8.2)	8.0 [7.7; 8.3]		2050 (8.3)	6.1 [5.8; 6.3]		11,260 (6.9)	40.2 [39.4; 40.9]		1008 (8.4)	2.9 [2.8; 3.1]		1650 (8.3)	4.9 [4.7; 5.1]	
65−69 (n = 29,315)	1532 (5.6)	9.3 [8.9; 9.8]		944 (5.7)	5.6 [5.3; 6.0]		6647 (4.9)	46.0 [45.0; 47.2]		411 (5.8)	2.4 [2.2; 2.7]		975 (5.7)	5.9 [5.5; 6.2]	
≥70 (n = 14,351)	522 (2.9)	12.6 [11.6; 13.8]		334 (2.9)	8.0 [7.2; 8.9]		2558 (2.6)	68.4 [65.8; 71.1]		77 (2.9)	1.8 [1.5; 2.3]		330 (2.9)	7.9 [7.1; 8.8]	
Breast density 1st mammogram			*			*			*			*			*
Type I (n = 35,046)	893 (7.2)	3.5 [3.3; 3.8]		1042 (7.2)	4.1 [3.9; 4.4]		8518 (6.1)	39.9 [39.1; 40.8]		543 (7.3)	2.1 [2.0; 2.3]		910 (7.2)	3.6 [3.4; 3.9]	
Type II (n = 161,248)	8273 (7.1)	7.3 [7.1; 7.4]		7080 (7.1)	6.1 [6.0; 6.3]		39,815 (6.0)	40.8 [40.4; 41.2]		4082 (7.2)	3.5 [3.4; 3.6]		5927 (7.1)	5.2 [5.0; 5.3]	
Type III (n = 58,520)	5199 (7.2)	12.3 [12.0; 12.7]		3056 (7.5)	7.0 [6.8; 7.2]		13,424 (6.4)	35.6 [35.0; 36.2]		1579 (7.6)	3.5 [3.4; 3.7]		2329 (7.5)	5.3 [5.1; 5.6]	
Type IV (n = 6040)	675 (7.2)	15.5 [14.4; 16.7]		271 (7.7)	5.8 [5.2; 6.6]		1175 (6.8)	28.5 [26.9; 30.2]		105 (7.8)	2.2 [1.8; 2.7]		166 (7.7)	3.6 [3.1; 4.2]	
HRT 1st mammogram			*			*			*			*			*
No/U (n = 227,111)	12,519 (6.9)	7.9 [7.8; 8.1]		9523 (7.0)	5.9 [5.8; 6.1]		53,780 (6.0)	39.4 [39.0; 39.7]		5403 (7.1)	3.3 [3.2; 3.4]		7895 (7.0)	5.0 [4.8; 5.1]	
Yes (n = 33,743)	2521 (8.4)	8.9 [8.6; 9.3]		1926 (8.5)	6.7 [6.4; 7.0]		9152 (7.1)	38.0 [37.2; 38.8]		906 (8.7)	3.1 [2.9; 3.3]		1437 (8.5)	5.0 [4.7; 5.3]	

*Log-rank test *P* value < 0.0001; Tp, average time at risk (in year).
[a] 258 women did not have the left breast.
[b] 229 women did not have the right breast − C-In [CI 95%], cumulative incidence per 1000 person-year [95% confidence interval]; P-y, person-year; No/U, No/unspecified; HRT, hormone replacement therapy.

Was considered second reading cancer (R2-Cancer) when the diagnostic process was started after a positive mammographic in the second reading session (ACR in 0,3,4,5) following a negative first reading (ACR in 1,2). Interval cancers were those detected in the interim between regular screening examinations.

The independent factors were (1) the abnormality discovery age (in five categories: *50–54, 55–59, 60–64, 65–69*, and *70–74* years); (2) the presence of the abnormality in the contralateral breast if it is present in the examined breast (in five categories: the absence of the abnormality in the examined breast '*NA*', the absence of the abnormality in the contralateral breast while it is present in the examined breast '*Absent*', the presence of the abnormality in the contralateral breast before its presence in the examined breast '*Before*', the presence of the abnormality in both breasts on the same date '*Same_date*', the presence of the abnormality in the contralateral breast after its presence in the examined breast '*After*'); (3) together (abnormalities discovered immediately on the same date) or sequential (abnormalities in successive discovery over time) discovery over time of ≥2 mammographic abnormalities in the examined breast (in six categories: no abnormality '*NA*', an isolated abnormality '*1-isolated*', two abnormalities discovered together '*2-together*', two abnormalities discovered sequentially '*2-sequential*', ≥3 abnormalities discovered together '*3-together*', ≥3 abnormalities discovered sequentially '*3-sequential*'); and (4) the abnormalities occurrence's order (in six categories: Spiculated mass first, microcalcification first, obscured mass first, asymmetric density first, architectural distortion first, ≥2 abnormalities discovered first); (5) use of hormone replacement therapy (HRT; yes, no, and uncertain); (vi) breast density (types I to IV).

Statistical analysis

Student's *t*-test was used to compare the age groups (censored vs others).

All cumulative incidences (abnormality or B-cancer) were estimated by the actuarial life table method. Their 95% confidence intervals were estimated by the Greenwood method.

In incidence analysis (each mammographic abnormality), the women contributed to the calculation of person-times starting from the date of the first mammogram until the date of discovery of the abnormality in the examined breast or until the end date. Similarly, in the B-cancer incidence analysis, women contributed to the calculation of person-times starting from the date of the first mammogram until the date of cancer occurrence in the examined breast or, in the absence of B-cancer, until the end date. The cumulative incidences were described and compared between groups using the confidence interval (CI) comparison and the log-rank test.

The B-cancer risk analysis according to the age at the discovery of the mammographic abnormalities was carried out by estimating the adjusted relative risk (RRa) using a multivariate Cox model. Only women who had at least one of the five abnormalities were included in this analysis. The Cox model included all covariates regardless of their *P*-values in univariate analysis. Because of their multiple collinearities, the five variables that describe the presence of the abnormalities in the contralateral breast were introduced into the final model in three modalities (NA, absent, and present ['Before'+'Same_date'+'After']). The parameters of the model were estimated using the maximum likelihood method. All analyses were performed using STATA software version 13 (College Station, TX, USA). The threshold of statistical significance was 5%.

Results

The study included 261,083 women whose mean age (±standard deviation) at first mammogram was 57.3 ± 6.6 years. Of these, 40,208 (15.4%) had a censored follow-up before 74 years of age because of death (132), loss to follow-up (38,194), relocation (411), or refusal to take part (1471). At baseline, women censured for refusal or death were significantly older than the others (59.2 ± 5.8 years vs 57.3 ± 6.6 years, $P < 0.0001$ and 58.7 ± 5.7 years vs 57.3 ± 6.6 years, $P < 0.0001$, respectively), but women censored for relocation or loss to follow-up were significantly younger (55.8 ± 5.0 years vs 57.3 ± 6.6 years, $P < 0.0001$ and 56.3 ± 5.0 years vs 57.3 ± 6.6 years, $P < 0.0001$, respectively).

Table 2
Breast cancer cumulative incidence according to the presence of mammographic abnormalities.

Mammographic abnormalities	Cumulative incidence									
	Left breast					Right breast				
	Nb of women (Tp)	P–Y	Case	C–In [CI 95%]	*P**	Nb of women (Tp)	P–Y	Case	C–In [CI 95%]	*P**
Overall	260,825 (7.6)[a]	1,953,560.6	3394	1.7 [1.7; 1.8]		260,854 (7.5)[b]	1,953,706.4	3059	1.6 [1.5; 1.6]	
Presence					$<10^{-3}$					$<10^{-3}$
No abnormality	173,233 (7.1)	1,229,252.5	785	0.6 [0.6; 0.7]		175,696 (7.1)	1,249,074.8	731	0.6 [0.5; 0.6]	
≥1 abnormality	87,592 (8.3)	724,308.1	2609	3.6 [3.5; 3.7]		85,158 (8.3)	704,631.6	2328	3.3 [3.2; 3.4]	
Microcalcification					$<10^{-3}$					$<10^{-3}$
No	245,568 (7.4)	1,826,457.0	2417	1.3 [1.3; 1.4]		245,814 (7.4)	1,827,982.0	2185	1.2 [1.1; 1.2]	
Yes	15,257 (8.3)	127,103.6	977	7.7 [7.2; 8.2]		15,040 (8.4)	125,724.5	874	7.0 [6.5; 7.4]	
Spiculated mass					$<10^{-3}$					$<10^{-3}$
No	248,809 (7.4)	1,848,301.9	2587	1.4 [1.3; 1.5]		249,405 (7.4)	1,853,980.0	2370	1.3 [1.2; 1.3]	
Yes	12,016 (8.9)	105,258.7	807	7.7 [7.2; 8.2]		11,449 (8.7)	99,726.5	689	6.9 [6.4; 7.4]	
Obscured mass					$<10^{-3}$					$<10^{-3}$
No	195,320 (7.2)	1,407691.3	1959	1.4 [1.3; 1.5]		197,922 (7.2)	1,429,130.8	1794	1.3 [1.2; 1.3]	
Yes	65,505 (8.3)	545,869.3	1435	2.6 [2.5; 2.8]		62,932 (8.3)	524,575.7	1265	2.4 [2.3; 2.5]	
Asymmetric density					$<10^{-3}$					$<10^{-3}$
No	254,460 (7.5)	1,900756.4	2983	1.6 [1.5; 1.6]		254,545 (7.5)	1,901,211.3	2715	1.4 [1.4; 1.5]	
Yes	6365 (8.3)	52,804.2	411	7.8 [7.1; 8.6]		6309 (8.3)	52,495.1	344	6.6 [5.9; 7.3]	
Architectural distortion					$<10^{-3}$					$<10^{-3}$
No	251,144 (7.4)	1,869595.7	3034	1.6 [1.6; 1.7]		251,522 (7.5)	1,872,887.6	2671	1.4 [1.4; 1.5]	
Yes	9681 (8.7)	83,964.9	360	4.3 [3.9; 4.8]		9332 (8.7)	80,818.9	388	4.8 [4.3; 5.3]	

Tp, average time at risk (in years).
*P value of the log-rank test.
[a] 258 women did not have the left breast.
[b] 229 women did not have the right breast – C–In [CI 95%], cumulative incidence per 1000 person-year [95% confidence interval]; Nb, number; P–Y, person-year.

Table 3
Cumulative incidence by type of cancer's localization (left/right), according to the characteristics of the radiological abnormalities and the characteristics of the women.

Characteristics	Left breast cancer's cumulative incidence				Right breast cancer's cumulative incidence			
	Nb of women	Case	C−In [CI 95%]	*P**	Nb of women	Case	C−In [CI 95%]	*P**
Overall	*260,825*[a]	3394	1.7 [1.7; 1.8]		*260,854*[b]	3059	1.6 [1.5; 1.6]	
Finding age(year): microcalcification				$<10^{-3}$				$<10^{-3}$
Absent	245,568	2417	1.3 [1.3; 1.4]		245,814	2185	1.2 [1.1; 1.2]	
50−54	4328	190	5.8 [5.1; 6.7]		4241	157	4.9 [4.2; 5.7]	
55−59	3544	214	6.7 [5.9; 7.7]		3469	181	5.8 [5.0; 6.7]	
60−64	3178	253	8.4 [7.4; 9.5]		3175	214	7.2 [6.3; 8.2]	
65−69	2458	174	8.4 [7.2; 9.7]		2458	177	8.5 [7.3; 9.8]	
≥70	1749	146	12.2 [10.4; 14.4]		1697	145	12.5 [10.6; 14.7]	
Contralateral presence if a microcalcification seen				$<10^{-3}$				$<10^{-3}$
Absent	5631	650	14.3 [13.2; 15.4]		5414	554	12.5 [11.5; 13.6]	
Before	287	17	6.2 [3.9; 10.0]		286	27	10.1 [6.9; 14.7]	
Same date	9053	295	3.9 [3.5; 4.3]		9053	285	3.7 [3.3; 4.2]	
After	286	15	5.6 [3.4; 9.3]		287	8	2.9 [1.5; 5.9]	
Finding age(year): spiculated mass				$<10^{-3}$				$<10^{-3}$
Absent	248,809	2587	1.4 [1.3; 1.5]		249,405	2370	1.3 [1.2; 1.3]	
50−54	2593	78	4.2 [3.3; 5.2]		2438	74	4.4 [3.5; 5.5]	
55−59	2813	151	6.0 [5.1; 7.1]		2687	118	4.9 [4.1; 5.9]	
60−64	2738	199	7.3 [6.3; 8.3]		2669	181	6.8 [5.8; 7.8]	
65−69	2278	215	10.0 [8.8; 11.4]		2195	180	8.7 [7.5; 10.1]	
≥70	1594	164	13.0 [11.2; 15.2]		1460	136	11.8 [10.0; 14.0]	
Contralateral presence if a spiculated mass seen				$<10^{-3}$				$<10^{-3}$
Absent	8581	753	10.1 [9.4; 10.9]		8014	642	9.3 [8.6; 10.1]	
Before	336	12	3.6 [2.0; 6.3]		348	10	2.8 [1.5; 5.2]	
Same date	2751	36	1.5 [1.1; 2.1]		2751	32	1.3 [0.9; 1.9]	
After	348	6	1.7 [0.8; 3.8]		336	5	1.5 [0.6; 3.6]	
Finding age(year): obscured mass				$<10^{-3}$				$<10^{-3}$
Absent	195,320	1959	1.4 [1.3; 1.5]		197,922	1794	1.3 [1.2; 1.3]	
50−54	18,256	291	2.2 [1.9; 2.4]		17,553	236	1.8 [1.6; 2.1]	
55−59	15,534	345	2.4 [2.2; 2.7]		14,874	298	2.2 [2.0; 2.5]	
60−64	14,120	373	2.7 [2.5; 3.0]		13,709	336	2.5 [2.3; 2.8]	
65−69	10,552	274	3.1 [2.7; 3.5]		10,030	253	3.0 [2.7; 3.4]	
≥70	7043	152	3.4 [2.9; 4.0]		6766	142	3.3 [2.8; 3.9]	
Contralateral presence if an obscured mass seen				$<10^{-3}$				$<10^{-3}$
Absent	32,813	891	3.4 [3.1; 3.6]		30,240	760	3.1 [2.9; 3.3]	
Before	5521	117	2.3 [1.9; 2.7]		5750	140	2.6 [2.2; 3.0]	
Same date	21,421	351	2.0 [1.8; 2.2]		21,421	313	1.8 [1.6; 2.0]	
After	5750	76	1.4 [1.1; 1.8]		5521	52	1.0 [0.8; 1.3]	
Finding age(year): asymmetric density				$<10^{-3}$				$<10^{-3}$
Absent	254,460	2983	1.6 [1.5; 1.6]		254,545	2715	1.4 [1.4; 1.5]	
50−54	1482	54	6.8 [5.2; 8.9]		1409	40	5.5 [4.0; 7.4]	
55−59	1415	79	6.7 [5.4; 8.4]		1422	84	7.1 [5.8; 8.8]	
60−64	1444	106	7.4 [6.2; 9.0]		1486	85	5.8 [4.7; 7.2]	
65−69	1201	93	7.9 [6.5; 9.7]		1219	79	6.6 [5.3; 8.2]	
≥70	823	79	11.2 [9.0; 13.9]		773	56	8.3 [6.4; 10.8]	
Contralateral presence if an asymmetry seen				$<10^{-3}$				$<10^{-3}$
Absent	5262	405	9.3 [8.4; 10.3]		5206	333	7.7 [6.9; 8.6]	
Before	162	2	1.4 [0.4; 5.6]		190	0	0	
Same date	751	4	0.7 [0.2; 1.7]		751	9	1.5 [0.8; 2.8]	
After	190	0	0		162	2	1.4 [0.4; 5.6]	
Finding age(year): architectural distortion				$<10^{-3}$				$<10^{-3}$
Absent	251,144	3034	1.6 [1.6; 1.7]		251,522	2671	1.4 [1.4; 1.5]	
50−54	2646	60	2.8 [2.2; 3.6]		2517	76	3.8 [3.1; 4.8]	
55−59	2415	71	3.1 [2.5; 3.9]		2329	72	3.3 [2.6; 4.1]	
60−64	2090	98	4.8 [3.9; 5.8]		2045	84	4.2 [3.4; 5.2]	
65−69	1561	70	5.2 [4.1; 6.6]		15	96	7.5 [6.1; 9.2]	
≥70	969	61	10.1 [7.9; 13.0]		941	60	10.2 [7.9; 13.1]	
Contralateral presence if a distortion seen				$<10^{-3}$				$<10^{-3}$
Absent	7043	328	5.4 [4.8; 6.0]		6694	358	6.2 [5.6; 6.9]	
Before	246	10	4.1 [2.2; 7.7]		258	5	2.0 [0.8; 4.7]	
Same date	2134	20	1.1 [0.7; 1.7]		2134	23	1.3 [0.9; 1.9]	
After	258	2	0.8 [0.2; 3.1]		246	2	0.8 [0.2; 3.3]	
Abnormalities discovery over time				$<10^{-3}$				$<10^{-3}$
No abnormality	173,233	785	0.6 [0.6; 0.7]		175,696	731	0.6 [0.5; 0.6]	
1-isolated	69,116	1517	2.7 [2.6; 2.8]		67,823	1358	2.5 [2.3; 2.6]	
2-together	6702	397	7.4 [6.7; 8.1]		6294	340	6.7 [6.0; 7.4]	
2-sequential	9222	444	5.1 [4.7; 5.6]		8685	416	5.1 [4.7; 5.7]	
3-together	413	67	20.3 [16.0; 25.8]		397	49	15.4 [11.7; 20.4]	
3-sequential	2139	184	8.8 [7.6; 10.2]		1959	165	8.6 [7.4; 10.0]	
Abnormalities occurrence's order				$<10^{-3}$				$<10^{-3}$
Spiculated mass first	6725	348	6.1 [5.5; 6.8]		6576	288	5.2 [4.6; 5.8]	
Obscured mass first	53,019	856	2.0 [1.8; 2.1]		51,183	776	1.8 [1.7; 2.0]	
Microcalcification first	7569	394	6.4 [5.8; 7.1]		7815	350	5.5 [4.9; 6.1]	

(*continued on next page*)

Table 3 (*continued*)

Characteristics	Left breast cancer's cumulative incidence				Right breast cancer's cumulative incidence			
	Nb of women	Case	C–In [CI 95%]	*P**	Nb of women	Case	C–In [CI 95%]	*P**
Architectural distortion first	7087	179	2.9 [2.5; 3.4]		6916	208	3.5 [3.0; 4.0]	
Asymmetric density first	4328	247	7.1 [6.3; 8.1]		4305	208	6.0 [5.3; 6.9]	
≥2 abnormalities first	8864	585	7.9 [7.3; 8.6]		8363	498	7.2 [6.6; 7.8]	
Hormone replacement therapy				$<10^{-3}$				$<10^{-3}$
No/unspecified	227,085	2785	1.7 [1.6; 1.7]		227,111	2545	1.5 [1.5; 1.6]	
Yes	33,740	609	2.0 [1.9; 2.2]		33,743	514	1.7 [1.6; 1.9]	
Breast density				$<10^{-3}$				$<10^{-3}$
D-I	35,040	322	1.2 [1.1; 1.4]		35,046	267	1.0 [0.9; 1.2]	
D-II	161,222	2059	1.7 [1.7; 1.8]		161,248	1876	1.6 [1.5; 1.6]	
D-III	58,524	926	2.0 [1.9; 2.2]		5852	827	1.8 [1.7; 2.0]	
D-IV	6039	87	1.8 [1.5; 2.2]		6040	89	1.9 [1.5; 2.3]	

*P value of the log-rank test.
[a] 258 did not have the left breast.
[b] 229 did not have the right breast – C-In [CI 95%]: cumulative incidence/1000P–Y [95% confidence interval]; Nb: Number.

Incidence of mammographic abnormalities

On average, women had 4.1 ± 1.7 mammograms, and the mean delay between two consecutive mammograms was 2.5 ± 0.9 years. At least one of the five abnormalities was seen in 122,343 women (46.9%) at an average age of 60.0 ± 6.8 years. The average duration of follow-up was 4.9 years. The overall incidence of abnormality (all abnormalities, both breast) was estimated at 95.4/1000 person-years (p-y; IC: 94.9; 95.9). A total of 87,592 women had at least one abnormality in the left breast (incidence 59.2/1000 p-y [58.8; 59.5]), and 85,158 women had at least one abnormality in the right breast (incidence 56.9/1000 p-y [56.6; 57.3]).

The mean age at discovery of the first abnormality was 60.3 ± 6.7 years (range: 51–74 years). The mean follow-up (time at risk) before this discovery ranged from 6.1 years (obscured mass) to 7.2 years (spiculated mass; Table 1). Regardless of laterality, the incidence of microcalcifications increased, whereas the incidence of obscured mass decreased with breast density.

In the left breast, the first abnormality detected at a mean age of 60.3 ± 6.8 years was microcalcification (8.6% of cases) or spiculated mass (7.7%), or obscured mass (60.5%), or asymmetric density (4.9%), or architectural distortion (7.1%), or a combination of at least two abnormalities (11.2%). In the right breast, the mean age at the discovery of the first abnormality was 60.2 ± 6.7 years, and the proportions of the abnormalities cited previously were, respectively, 9.2%, 7.7%, 60.1%, 5.1%, 8.1%, and 9.8%. In the left breast, this first abnormality was discovered in 40.8% of cases at the first mammogram (M1), 24.3% at M2, 14.8% at M3, and 20.1% at M4 or later. In the right breast, these percentages were, respectively, 40.5, 24.4, 15.0, and 20.2 at M4 or later.

Incidence of breast cancer

The mean follow-up time (261,083 women) was 7.4 years. At least one B-cancer was diagnosed in 6326 women (3137 left breast, 2802 right breast, 257 bilateral, and 130 unspecified location), which represents an incidence of 3.3/1000 p-y (3.0; 3.1). The B-cancers were ADK-I (63.8%), TIS (10.5%), 'rare form' (2.7%) or unspecified (23.0%). These B-cancers were R1-Cancer (84.5%), R2-Cancer (4.5%) or interval cancers (11.0%). Overall, 88.4% of the 130 B-cancers whose laterality (left/right) was not specified were interval cancers.

The mean age at diagnosis was 63.8 ± 6.3 years, and 20.5% of B-cancers were diagnosed in women with none of the five mammographic abnormalities. Among 2414 cases of the 5028 cases associated with at least one mammographic abnormality, the abnormality (single in 64.4% of cases) was seen on the mammography that initiated the B-cancer diagnostic procedure.

B-cancer cumulative incidence was significantly lower in women with no mammographic abnormality than in those with at least one abnormality (right breast: 0.6 [0.5; 0.6] vs 3.3 [3.2; 3.4]). The incidence of B-cancer was five times higher in women with than without spiculated mass (right breast: 6.9 [6.4; 7.4] vs 1.3 [1.2; 1.3]; Table 2).

Although the confidence intervals are overlapped a few times, the log-rank test concludes that there is a significant difference regardless of the explanatory variable (Table 3). The incidence of B-cancer increased significantly ($P < 0.0001$) with the age at discovery of microcalcification (left breast: 50–54 years: 5.8 [5.1; 6.7]; ≥70 years: 12.2 [10.4; 14.4]). Whatever the abnormality, the incidence of B-cancer was higher when the abnormality was present only in the examined breast (absent in the contralateral breast; Table 3).

The incidence was constant between the second and eighth mammograms. Women with spiculated mass discovered before any other abnormality had a higher incidence between M2 and M7 than other women (Fig. 1A). Similarly, women in whom at least three abnormalities were discovered together or sequentially had a higher incidence rate between M2 and M7 (Fig. 1-B). Compared with 50–54 years, the B-cancer risk was 1.4 times higher when the architectural distortion discovery age was 65–69 years (right breast: RR_a: 1.4 [1.1; 2.0], $P = 0.02$; Table 4).

Discussion

This study, which involved, on average, a series of four mammograms per woman, showed a high incidence of B-cancer based on five radiological abnormalities. The B-cancer risk increased with the age at discovery of spiculated mass and microcalcification. Whatever the abnormality, the B-cancer risk was higher when the abnormality was present only in the examined breast. The study made it possible to estimate the mean delays(D) between three dates: (1) D1: 'first mammogram' to 'radiological abnormality discovery'; (2) D2: 'first mammogram' to 'B-cancer diagnosis'; (3) D3: 'radiological abnormality discovery' to 'B-cancer diagnosis'. When the mammographic finding was a spiculated mass, an asymmetric density, or an architectural distortion, D3 was shorter than the usual 2-year delay between two mammograms in the BCSP. This strengthens the recommendation to perform other mammograms between two campaigns in case of mammographic finding classified ACR3.

Although the observed delay between two consecutive mammograms was greater than the usual 2 years,[2–4] this study showed that having a normal mammogram does not reduce the B-cancer risk during the next mammogram. Actually, this stability between the second and the seventh mammograms is consistent with trends in the stability of incidence already described in France[30] and Spain.[31]

The increased risk of B-cancer with the spiculated mass discovery age observed here is comparable to a synergy between two risk factors (age and spiculated mass). Indeed, the link between age and B-cancer and the link between B-cancer and mammographic abnormality has been described.[6,9–11,13,32] In addition, the lower risk associated with radiological abnormalities in both breasts in comparison with an isolated abnormality in one breast is poorly documented. This should alert radiologists to the relevance of

Table 4
Breast cancer risk analysis in univariate and multivariate Cox model.

Characteristics	Left breast			Right breast		
	RR_u [CI 95%]	RR_a [CI 95%]	$P^{\dagger}$	RR_u [CI 95%]	RR_a [CI 95%]	$P^{\dagger}$
*Age(year) of diagnosis: microcalcifications**						
55–59 years	1.0 [0.9; 1.2]	1.0 [0.8; 1.2]	0.83	1.1 [0.9; 1.3]	1.1 [0.9; 1.4]	0.75
60–64 years	1.2 [1.1; 1.4]	1.1 [0.9; 1.4]	0.25	1.3 [1.0; 1.6]	1.2 [1.0; 1.5]	0.05
65–69 years	1.3 [1.1; 1.6]	1.0 [0.8; 1.3]	0.66	1.6 [1.3; 2.0]	1.3 [1.1; 1.7]	0.01
≥70 years	2.1 [1.7; 2.6]	1.4 [1.1; 1.7]	0.009	2.5 [2.0; 3.2]	1.9 [1.5; 2.4]	$<10^{-3}$
No microcalcifications	0.4 [0.4; 0.5]	0.3 [0.2; 0.4]	$<10^{-3}$	0.5 [0.4; 0.6]	0.2 [0.1; 0.3]	$<10^{-3}$
*Contralateral presence if a microcalcification seen***						
Present	0.3 [0.2; 0.3]	0.3 [0.2; 0.3]	$<10^{-3}$	0.3 [0.3; 0.4]	0.3 [0.3; 0.4]	$<10^{-3}$
NA	0.2 [0.2; 0.2]	0.3 [0.2; 0.4]	$<10^{-3}$	0.2 [0.2; 0.2]	0.2 [0.1; 0.3]	$<10^{-3}$
*Age(year) of diagnosis: spiculated mass**						
55–59 years	1.3 [1.0; 1.7]	1.4 [1.0; 1.8]	0.02	1.0 [0.8; 1.3]	1.1 [0.9; 1.4]	0.71
60–64 years	1.5 [1.1; 1.9]	1.4 [1.1; 1.9]	0.009	1.3 [1.0; 1.7]	1.3 [1.0; 1.7]	0.03
65–69 years	2.0 [1.6; 2.6]	2.0 [1.6; 2.6]	$<10^{-3}$	1.7 [1.3; 2.2]	1.6 [1.2; 2.1]	$<10^{-3}$
≥70 years	2.9 [2.2; 3.7]	2.4 [1.9; 3.3]	$<10^{-3}$	2.5 [1.9; 3.3]	2.0 [1.5; 2.7]	$<10^{-3}$
No spiculated mass	0.6 [0.5; 0.8]	0.7 [0.5; 1.0]	0.04	0.5 [0.4; 0.7]	0.3 [0.2; 0.5]	$<10^{-3}$
*Contralateral presence if a spiculated mass seen***						
Present	0.2 [0.1; 0.2]	0.3 [0.2; 0.6]	$<10^{-3}$	0.2 [0.1; 0.2]	0.2 [0.1; 0.2]	$<10^{-3}$
NA	0.3 [0.3; 0.3]	0.7 [0.5; 1.0]	0.04	0.3 [0.3; 0.3]	0.3 [0.2; 0.5]	$<10^{-3}$
*Age(year) of diagnosis: obscured mass **						
55–59 years	1.0 [0.8; 1.1]	1.0 [0.8; 1.1]	0.47	1.1 [0.9; 1.3]	1.0 [0.9; 1.2]	0.60
60–64 years	1.1 [0.9; 1.2]	1.0 [0.8; 1.2]	0.91	1.2 [1.0; 1.4]	1.1 [0.9; 1.3]	0.20
65–69 years	1.3 [1.1; 1.5]	1.2 [1.0; 1.4]	0.06	1.5 [1.2; 1.8]	1.3 [1.1; 1.6]	0.002
≥70 years	1.6 [1.3; 1.9]	1.4 [1.1; 1.7]	0.001	1.8 [1.5; 2.3]	1.6 [1.5; 2.3]	$<10^{-3}$
No obscured mass	2.8 [2.4; 3.2]	1.8 [1.3; 2.6]	0.001	3.0 [2.6; 3.4]	1.2 [0.6; 0.8]	0.33
*Contralateral presence if an obscured mass seen***						
Present	0.6 [0.5; 0.6]	0.6 [0.5; 0.8]	$<10^{-3}$	0.6 [0.5; 0.6]	0.6 [0.5; 0.6]	$<10^{-3}$
NA	2.0 [1.8; 2.1]	1.8 [1.3; 2.6]	0.001	1.9 [1.7; 2.1]	1.2 [0.6; 0.8]	0.33
*Age(year) of diagnosis: asymmetric density **						
55–59 years	0.8 [0.5; 1.1]	0.8 [0.5; 1.1]	0.12	1.0 [0.7; 1.5]	0.9 [0.6; 1.4]	0.69
60–64 years	0.7 [0.5; 1.0]	0.7 [0.5; 0.9]	0.01	0.7 [0.5; 1.0]	0.7 [0.5; 1.0]	0.05
65–69 years	0.7 [0.5; 1.0]	0.7 [0.5; 1.0]	0.07	0.8 [0.5; 1.1]	0.7 [0.5; 1.0]	0.06
≥70 years	1.1 [0.8; 1.6]	0.9 [0.6; 1.3]	0.47	1.1 [0.7; 1.6]	0.8 [0.5; 1.2]	0.33
No *asymmetric density*	0.4 [0.3; 0.5]	0.4 [0.3; 0.7]	$<10^{-3}$	0.4 [0.3; 0.6]	0.3 [0.2; 0.5]	$<10^{-3}$
*Contralateral presence if an asymmetry seen***						
Present	0.1 [0.0; 0.1]	0.1 [0.0; 0.2]	$<10^{-3}$	0.1 [0.1; 0.3]	0.2 [0.1; 0.3]	$<10^{-3}$
NA	0.4 [0.3; 0.4]	0.4 [0.3; 0.7]	$<10^{-3}$	0.4 [0.4; 0.5]	0.3 [0.2; 0.5]	$<10^{-3}$
*Age(year) of diagnosis: architectural distortion **						
55–59 years	1.0 [0.7; 1.3]	1.0 [0.7; 1.4]	0.95	0.7 [0.5; 1.0]	0.8 [0.6; 1.1]	0.11
60–64 years	1.4 [1.0; 2.0]	1.3 [0.9; 1.8]	0.10	0.9 [0.7; 1.3]	0.9 [0.7; 1.3]	0.58
65–69 years	1.7 [1.2; 2.4]	1.4 [1.0; 2.0]	0.03	1.8 [1.3; 2.4]	1.4 [1.1; 2.0]	0.02
≥70 years	3.9 [2.7; 5.5]	2.5 [1.7; 3.5]	$<10^{-3}$	2.7 [2.0; 3.9]	1.7 [1.2; 2.5]	0.002
No distortion	1.2 [0.9; 1.5]	1.0 [0.7; 1.6]	0.84	0.8 [0.6; 1.0]	0.4 [0.3; 0.6]	$<10^{-3}$
*Contralateral presence if a distortion seen***						
Present	0.3 [0.2; 0.4]	0.3 [0.2; 0.5]	$<10^{-3}$	0.2 [0.1; 0.3]	0.3 [0.2; 0.4]	$<10^{-3}$
NA	0.7 [0.6; 0.8]	1.0 [0.7; 1.6]	0.84	0.5 [0.5; 0.6]	0.4 [0.3; 0.6]	$<10^{-3}$
Abnormalities discovery over time (Ref.: 1-isolated)						
2-together	2.7 [2.4; 3.0]	2.0 [1.3; 3.1]	0.002	2.7 [2.4; 3.1]	2.2 [1.4; 2.9]	0.001
2-sequential	1.8 [1.6; 2.0]	1.6 [1.2; 2.3]	0.009	2.0 [1.8; 2.2]	2.0 [1.5; 2.6]	0.004
3-together	7.4 [5.8; 9.5]	3.0 [1.3; 6.7]	0.008	6.4 [4.8; 8.5]	2.4 [1.8; 3.1]	$<10^{-3}$
3-sequential	3.0 [2.6; 3.5]	2.0 [1.1; 4.0]	0.04	3.2 [2.7; 3.8]	3.3 [2.8; 3.8]	$<10^{-3}$
Abnormalities occurrence's order (Ref.: spiculated first)						
Obscured mass first	0.3 [0.3; 0.4]	1.0 [0.8; 1.2]	0.88	0.4 [0.3; 0.4]	1.2 [1.0; 1.5]	0.07
Microcalcification first	1.1 [0.9; 1.2]	0.9 [0.7; 1.1]	0.14	1.1 [0.9; 1.2]	0.9 [0.7; 1.1]	0.17
Architectural distortion first	0.5 [0.4; 0.5]	0.6 [0.5; 0.8]	0.002	0.6 [0.5; 0.8]	0.8 [0.6; 1.0]	0.07
Asymmetric density first	1.1 [0.9; 1.3]	1.2 [0.9; 1.6]	0.33	1.1 [0.9; 1.3]	1.4 [1.1; 1.9]	0.01
≥2 abnormalities first	1.3 [1.1; 1.5]	0.9 [0.7; 1.2]	0.48	1.3 [1.2; 1.6]	0.9 [0.7; 1.3]	0.63
Hormone replacement therapy (Ref.: No)						
Yes	1.1 [1.0; 1.2]	1.0 [0.9; 1.1]	0.53	1.0 [0.9; 1.1]	0.9 [0.8; 10]	0.07
Breast density (Ref.: Type I)						
Type II	1.2 [1.1; 1.4]	1.1 [1.0; 1.2]	0.21	1.3 [1.1; 1.5]	1.1 [1.0; 1.3]	0.06
Type III	1.3 [1.1; 1.5]	1.1 [1.0; 1.3]	0.13	1.4 [1.2; 1.6]	1.2 [1.0; 1.3]	0.08
Type IV	1.2 [0.9; 1.6]	1.0 [0.7; 1.3]	0.80	1.5 [1.1; 1.9]	1.3 [1.0; 1.7]	0.10

NA, absence of the abnormality in the examined breast; Ref., reference; RR_u, unadjusted relative risk from a univariate Cox model; RR_a, adjusted relative risk from a multivariate Cox model; *(Reference = 50–54years); **(Reference = 'Absent': the presence of the abnormality in the examined breast and absence of abnormality in the contralateral breast); [CI 95%], 95% confidence interval.[†]*P* value ($P > |z|$) from multivariate Cox model. Only women (87,592 women in left breast cancer risk analysis and 85,158 women in right breast cancer risk analysis) who had at least one of the five mammographic abnormalities were included in this multivariate risk analysis.

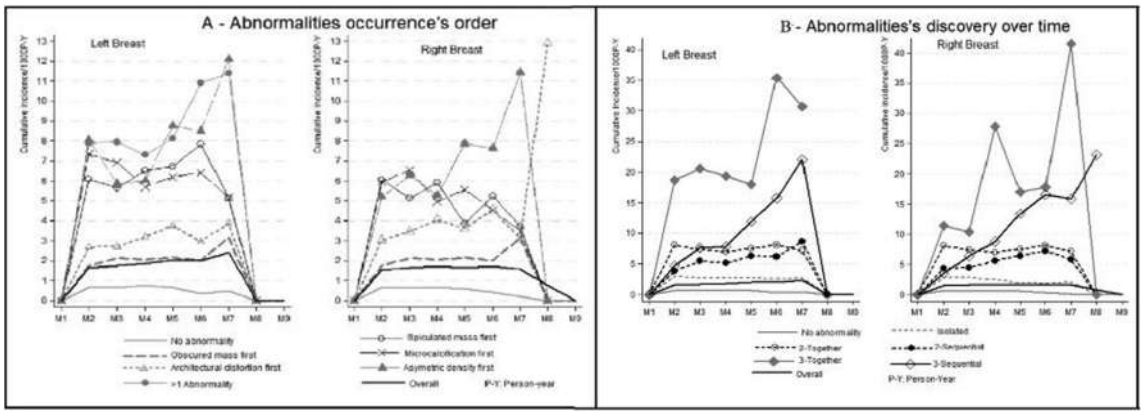


Fig. 1. **(A)** Evolution of the cumulative incidence according to the order of occurrence of the radiological abnormalities in the left and right breast and according by the screening mammogram (M) rank. No abnormality (no abnormality in the breast concerned); >1 abnormality (several abnormalities first occurred on the same date); overall (sample cumulative incidence). **(B)** Evolution of cumulative incidence according to the radiological abnormalities discovery over time in the left and right breasts and according to the rank of the screening mammogram (M). No abnormality (if no abnormality in the breast concerned); 1, isolated (1 isolated abnormality); 2, together (2 abnormalities discovered together); 2, sequential (2 abnormalities discovered sequentially); 3, together (≥3 abnormalities discovered together); 3, sequential (≥3 discovered sequentially); overall (sample cumulative incidence).

certain diagnostic procedures in the presence of a bilateral radiological finding, especially because this study showed that there is more localization of cancer in the left breast compared with the right breast. This lower risk would explain the low proportion of bilateral cancers in this study. Moreover, it has been shown that there is no apparent increase in the risk of developing a contralateral B-cancer according to the histology of primary cancer.[19,20] A quantitative analysis of homolateral views of mammograms would provide useful information regarding B-cancer risk over the short term.[16–18]

The absence of abnormality was related to the histopathological characteristics of the tumor modulated by patient's specific factors. Thus, a small-sized tumor, the absence of microcalcifications (often linked with tumor necrosis), or a minimal or absent stroma reaction does not facilitate lesion detection, especially in a radiologically dense breast.[14] In this study, any risk of B-cancer was not related to breast density.

By increasing breast density, HRT is associated with an increased risk of disagreement between mammogram readers, especially regarding breast for mass.[33] HRT is also associated with B-cancer.[31,34] In this study, women with HRT had a lower incidence of microcalcifications, a higher incidence of asymmetry, a higher incidence of B-cancer but same order risk of B-cancer.

In France, B-cancer's detection rate was stable since 2004;[35] it was estimated at 7/1000 p-y in a biennial BCSP campaign. In the current controversy over the usefulness of BCSP,[36–38] the high incidence showed in this study highlights the benefit of the BCSP and the need to strengthen the follow-up after the finding of a radiological abnormality. In terms of B-cancer morbidity estimation, the present results agree with other incidence studies.[12,31,39]

Limitations

The incidence rates found here may be underestimated because of incomplete data because of censoring before the age of 74 years for loss to follow-up or relocation while the results of last mammograms were classified ACR 3, 4 or 5.

Based on mammographic reading reports, this study cannot establish, with certainty, the link between a radiological abnormality seen during a mammogram and a similar abnormality seen during a subsequent mammogram.

The lack of interconnection between departmental databases does not make it possible to know the antecedents of women who have relocated one or more times between the age of 50–74 years. Women excluded because having only one participation in BCSP probably have a history of BCSP campaigns participation in other departments.

Conclusion

The study highlights a stable incidence of B-cancer between successive mammograms, an increased risk of B-cancer with the finding age of spiculated mass and microcalcification. The reduced delay between the abnormalities discovery date and the B-cancer diagnosis date would justify a specific follow-up protocol after the finding of certain mammographic abnormalities, in particular, the spiculated mass. The study highlights the low risk related to the presence of the same mammographic abnormality in both breasts compared with the presence of the isolated mammographic abnormality in one of the breasts. In this period, when the quality of the program remains compromised because of overdiagnostics and other diagnostic explorations deemed unnecessary by BCSP critics, these results should alert radiologists to the relevance of certain diagnostic procedures in the management of a bilateral mammographic abnormality.

Author statements

Acknowledgements

The authors are grateful to their colleagues from the BCSP management structures. The authors are grateful to Jean Iwaz (CHU de Lyon) and Claude Vernier (OECD Paris) for the thorough editing of the article.

Ethical approval

The data were anonymized before analysis. The BCSPs' databases are agreed by the French "Commission Nationale de l'Informatique et des Libertes (CNIL)."[40] Following the French current legislation, no ethical approval needed.

Funding

No financial or other relationships that could appear to have influenced the study.

Competing interests

None declared.

Authors' contributions

The study was conceived and designed by K.A. The data were acquired and collated by B.C., R.R., and C.S. and analyzed by K.A. and B.C. The study was drafted and revised critically by all authors (B.C., K.A., R.R., C.S., R.M.C., and S.N.), and S.N. was the study's guarantor. All authors gave final approval of the version to be published and have contributed to the study.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca - Cancer J Clin* 2018. https://doi.org/10.3322/caac.21492. PubMed PMID: 30207593.
2. Seradour B, Ancelle-Park R. Breast cancer screening: are results of French and international programmes comparable? *J Radiol* 2006;**87**(9):1009–14. PubMed PMID: 16936624.
3. JORF. *Arrêté du 29 septembre 2006 relatif aux programmes de dépistage des cancers. Journal Officiel de la République Française (JORF) n°295 du 21 décembre 2006 page 19240 texte n° 49 NOR: SANP0623877A*. 2006. Avalaible at: https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000000460656. [Accessed 7 June 2021].
4. Schoch C. Mass breast cancer screening program: a public health priority in France. *Bull Cancer* 2002;**89**(12):1079–80. PubMed PMID: 12525368.
5. No Authors listed. Recommendations for radiologists involved with the breast cancer screening program. *J Radiol* 2003;**84**(12 Pt 1):1921–32. PubMed PMID: 14710042.
6. Berment H, Becette V, Mohallem M, Ferreira F, Cherel P. Masses in mammography: what are the underlying anatomopathological lesions? *Diagn Interv Imaging* 2014;**95**(2):124–33. https://doi.org/10.1016/j.diii.2013.12.010. PubMed PMID: 24456895.
7. Eriksson M, Czene K, Pawitan Y, Leifland K, Darabi H, Hall P. A clinical model for identifying the short-term risk of breast cancer. *Breast Cancer Res* 2017;**19**(1):29. https://doi.org/10.1186/s13058-017-0820-y. PubMed PMID: 28288659.
8. Rominger MB, Sax EV, Figiel JH, Timmesfeld N. Occurrence and positive predictive value of additional nonmass findings for risk stratification of breast microcalcifications in mammography. *Can Assoc Radiol J* 2013;**64**(4):333–8. https://doi.org/10.1016/j.carj.2012.07.004. PubMed PMID: 23298860.
9. Boyer B, Russ E. Anatomical-radiological correlations: architectural distortions. *Diagn Interv Imaging* 2014;**95**(2):134–40. https://doi.org/10.1016/j.diii.2014.01.003. PubMed PMID: 24529951.
10. Henrot P, Leroux A, Barlier C, Genin P. Breast microcalcifications: the lesions in anatomical pathology. *Diagn Interv Imaging* 2014;**95**(2):141–52. https://doi.org/10.1016/j.diii.2013.12.011. PubMed PMID: 24525087.
11. Farshid G, Sullivan T, Downey P, Gill PG, Pieterse S. Independent predictors of breast malignancy in screen-detected microcalcifications: biopsy results in 2545 cases. *Br J Cancer* 2011;**105**(11):1669–75. https://doi.org/10.1038/bjc.2011.466. PubMed PMID: 22052156.
12. Nederend J, Duijm LE, Voogd AC, Groenewoud JH, Jansen FH, Louwman MW. Trends in incidence and detection of advanced breast cancer at biennial screening mammography in The Netherlands: a population based study. *Breast Cancer Res* 2012;**14**(1):R10. https://doi.org/10.1186/bcr3091. PubMed PMID: 22230363.
13. Barreau B, de Mascarel I, Feuga C, MacGrogan G, Dilhuydy MH, Picot V, et al. Mammography of ductal carcinoma in situ of the breast: review of 909 cases with radiographic-pathologic correlations. *Eur J Radiol* 2005;**54**(1):55–61. https://doi.org/10.1016/j.ejrad.2004.11.019. PubMed PMID: 15797293.
14. Roger P, Delfour C, Ragu N, Serre I, Baldet P, Taourel P. Carcinoma with false negative mammogram: when and how? *J Radiol* 2004;**85**(12 Pt 2):2063–7. PubMed PMID: 15692420.
15. Broeders MJ, Onland-Moret NC, Rijken HJ, Hendriks JH, Verbeek AL, Holland R. Use of previous screening mammograms to identify features indicating cases that would have a possible gain in prognosis following earlier detection. *Eur J Cancer* 2003;**39**(12):1770–5. PubMed PMID: 12888373.
16. Sun W, Tseng TB, Qian W, Saltzstein EC, Zheng B, Yu H, et al. A new near-term breast cancer risk prediction scheme based on the quantitative analysis of ipsilateral view mammograms. *Comput Methods Progr Biomed* 2018;**155**:29–38. https://doi.org/10.1016/j.cmpb.2017.11.019. PubMed PMID: 29512502.
17. Tan M, Pu J, Cheng S, Liu H, Zheng B. Assessment of a four-view mammographic image feature based fusion model to predict near-term breast cancer risk. *Ann Biomed Eng* 2015;**43**(10):2416–28. https://doi.org/10.1007/s10439-015-1316-5. PubMed PMID: 25851469.
18. Li Y, Fan M, Cheng H, Zhang P, Zheng B, Li L. Assessment of global and local region-based bilateral mammographic feature asymmetry to predict short-term breast cancer risk. *Phys Med Biol* 2018;**63**(2):025004. https://doi.org/10.1088/1361-6560/aaa096. PubMed PMID: 29226849.
19. Langlands F, White J, Kearins O, Cheung S, Burns R, Horgan K, et al. Contralateral breast cancer: incidence according to ductal or lobular phenotype of the primary. *Clin Radiol* 2016;**71**(2):159–63. https://doi.org/10.1016/j.crad.2015.10.030. PubMed PMID: 26703116.
20. de Glas NA, Engels CC, Bastiaannet E, van de Water W, Siesling S, de Craen AJ, et al. Contralateral breast cancer risk in relation to tumor morphology and age-in which patients is preoperative MRI justified? *Breast Cancer Res Treat* 2015;**150**(1):191–8. https://doi.org/10.1007/s10549-015-3294-6. PubMed PMID: 25677741.
21. Bennani-Baiti B, Baltzer PAT. [Artificial intelligence in the diagnosis of breast cancer : yesterday, today and tomorrow]. *Radiologe* 2019. https://doi.org/10.1007/s00117-019-00615-y. PubMed PMID: 31811325.
22. Watanabe AT, Lim V, Vu HX, Chim R, Weise E, Liu J, et al. Improved cancer detection using artificial intelligence: a retrospective evaluation of missed cancers on mammography. *J Digit Imag* 2019;**32**(4):625–37. https://doi.org/10.1007/s10278-019-00192-5. PubMed PMID: 31011956.
23. Geras KJ, Mann RM, Moy L. Artificial intelligence for mammography and digital breast tomosynthesis: current concepts and future perspectives. *Radiology* 2019;**293**(2):246–59. https://doi.org/10.1148/radiol.2019182627. PubMed PMID: 31549948.
24. Houssami N, Kirkpatrick-Jones G, Noguchi N, Lee CI. Artificial Intelligence (AI) for the early detection of breast cancer: a scoping review to assess AI's potential in breast screening practice. *Expet Rev Med Dev* 2019;**16**(5):351–62. https://doi.org/10.1080/17434440.2019.1610387. PubMed PMID: 30999781.
25. Rodriguez-Ruiz A, Lang K, Gubern-Merida A, Teuwen J, Broeders M, Gennaro G, et al. Can we reduce the workload of mammographic screening by automatic identification of normal exams with artificial intelligence? A feasibility study. *Eur Radiol* 2019;**29**(9):4825–32. https://doi.org/10.1007/s00330-019-06186-9. PubMed PMID: 30993432.
26. Thomassin-Naggara I, Balleyguier C, Ceugnart L, Heid P, Lenczner G, Maire A, et al. Artificial intelligence and breast screening: French Radiology Community position paper. *Diagn Interv Imaging* 2019;**100**(10):553–66. https://doi.org/10.1016/j.diii.2019.08.005. PubMed PMID: 31523026.
27. Cai H, Huang Q, Rong W, Song Y, Li J, Wang J, et al. Breast microcalcification diagnosis using deep convolutional neural network from digital mammograms. *Comput Math Methods Med* 2019;**2019**:2717454. https://doi.org/10.1155/2019/2717454. PubMed PMID: 30944574.
28. Carter SM, Rogers W, Win KT, Frazer H, Richards B, Houssami N. The ethical, legal and social implications of using artificial intelligence systems in breast cancer care. *Breast* 2019;**49**:25–32. https://doi.org/10.1016/j.breast.2019.10.001. PubMed PMID: 31677530.
29. WHO. *International classification of diseases*. Geneve: World Health Organization; 1990.
30. Molinie F, Billon-Delacour S, Allioux C, Blais S, Bercelli P, Lombrail P. [Incidence and prognostic factors of breast cancer diagnosed within and outside of the organised screening program in the department of Loire-Atlantique (1991-2002)]. *Rev Epidemiol Sante Publique* 2008;**56**(1):41–9. https://doi.org/10.1016/j.respe.2008.01.001. PubMed PMID: 18295424.
31. Roman M, Rue M, Sala M, Ascunce N, Bare M, Baroja A, et al. Trends in detection of invasive cancer and ductal carcinoma in situ at biennial screening mammography in Spain: a retrospective cohort study. *PLoS One* 2013;**8**(12):e83121. https://doi.org/10.1371/journal.pone.0083121. PubMed PMID: 24376649.
32. Kopans DB, Swann CA, White G, McCarthy KA, Hall DA, Belmonte SJ, et al. Asymmetric breast tissue. *Radiology* 1989;**171**(3):639–43. https://doi.org/10.1148/radiology.171.3.2541463. PubMed PMID: 2541463.
33. Meye NE, Schaffer P, Hedelin G, Guldenfels C, Gairard B. Impact of menopause hormone replacement therapy on screening mammography reading. *J Radiol* 2001;**82**(6 Pt 1):653–9. PubMed PMID: 11449167.
34. Cordina-Duverger E, Truong T, Anger A, Sanchez M, Arveux P, Kerbrat P, et al. Risk of breast cancer by type of menopausal hormone therapy: a case-control study among post-menopausal women in France. *PLoS One* 2013;**8**(11):e78016. https://doi.org/10.1371/journal.pone.0078016. PubMed PMID: 24223752.
35. Quintin C, Rogel A. *Évaluation du programme de dépistage organisé du cancer du sein : résultats et évolution des indicateurs de performance depuis 2004 en France métropolitaine*. Saint-Maurice, France: Santé publique France; 2019.
36. Launoy G, Duchange N, Darquy S, Moutel G. Participation to organized screening programs: individual and collective stakes. *Bull Cancer* 2018. https://doi.org/10.1016/j.bulcan.2018.04.008. PubMed PMID: 29853200.
37. Pivot X, Viguier J, Touboul C, Morere JF, Blay JY, Coscas Y, et al. Breast cancer screening controversy: too much or not enough? *Eur J Cancer Prev* 2015;**24**(Suppl):S73–6. https://doi.org/10.1097/CEJ.0000000000000145. PubMed PMID: 26016791.
38. Perrenoud C, Stiefel F, Bourquin C. From facts to arguments: a study of the 2014 Swiss controversy over systematic mammography screening. *Patient Educ Counsel* 2018;**101**(6):1110–5. https://doi.org/10.1016/j.pec.2017.12.015. PubMed PMID: 29305062.
39. Natal C, Caicoya M, Prieto M, Tardon A. Breast cancer incidence related with a population-based screening program. *Med Clin (Barc).* 2015;**144**(4):156–60. https://doi.org/10.1016/j.medcli.2014.04.028. PubMed PMID: 25194975.

40. JORF. *Délibération n° 2017-215 du 13 juillet 2017 portant adoption d'une norme destinée à simplifier l'obligation de déclaration des traitements de données à caractère personnel ayant pour finalité le dépistage organisé du cancer du sein, du cancer colorectal et du cancer du col de l'utérus mis en œuvre par les structures de gestion conventionnées, et abrogeant la délibération n° 2015-175 du 11 juin 2015 (décision d'autorisation unique n° AU-043) (NS-059). Journal Officiel de la République Française (JORF) du 1er septembre 2017; n°0204 Texte n° 67. NOR: CNIL1724568X*. Avalaible at: https://www.legifrance.gouv.fr/jorf/article_jo/JORFARTI000035484855. [Accessed 7 June 2021].

Public Health 202 (2022) 84–92

Contents lists available at ScienceDirect

Public Health

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

Original Research

Clinical features, risk factors and a prediction model for in-hospital mortality among diabetic patients infected with COVID-19: data from a referral centre in Iran

Maryam Kabootari [a, b], Reza Habibi Tirtashi [a], Mitra Hasheminia [c], Mohammadreza Bozorgmanesh [d], Davood Khalili [b, c], Hamideh Akbari [a], Gholamreza Roshandel [e], Farzad Hadaegh [b, *]

[a] Metabolic Disorders Research Center, Golestan University of Medical Sciences, Gorgan, Iran
[b] Prevention of Metabolic Disorders Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[c] Department of Biostatistics and Epidemiology, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[d] Department of Orthopedic Surgery, Arak University of Medical Sciences, Arak, Iran
[e] Golestan Research Center of Gastroenterology and Hepatology, Golestan University of Medical Sciences, Gorgan, Iran

ARTICLE INFO

Article history:
Received 16 May 2021
Received in revised form
18 October 2021
Accepted 9 November 2021
Available online 17 November 2021

Keywords:
COVID-19
Diabetes
In-hospital mortality

ABSTRACT

Objectives: The aim of this study was to identify risk factors of in-hospital mortality among diabetic patients infected with COVID-19.
Study design: This is a retrospective cohort study.
Methods: Using logistic regression analysis, the independent association of potential prognostic factors and COVID-19 in-hospital mortality was investigated in three models. Model 1 included demographic data and patient history; model 2 consisted of model 1, plus vital signs and pulse oximetry measurements at hospital admission; and model 3 included model 2, plus laboratory test results at hospital admission. The odds ratios (ORs) and 95% confidence intervals (95% CIs) were reported for each predictor in the different models. Moreover, to examine the discriminatory powers of the models, a corrected area under the receiver-operating characteristic curve (AUC) was calculated.
Results: Among 560 patients with diabetes (men = 291) who were hospitalised for COVID-19, the mean age of the study population was 61.8 (standard deviation [SD] 13.4) years. During a median length of hospitalisation of 6 days, 165 deaths (men = 93) were recorded. In model 1, age and a history of cognitive impairment were associated with higher mortality; however, taking statins, oral antidiabetic drugs and beta-blockers was associated with a lower risk of mortality (AUC = 0.76). In model 2, adding the data for respiratory rate (OR 1.07 [95% CI 1.00–1.14]) and oxygen saturation (OR 0.95 [95% CI 0.92–0.98]) slightly increased the AUC to 0.80. In model 3, the data for platelet count (OR 0.99 [95% CI 0.99–1.00]), lactate dehydrogenase (OR 1.002 [95% CI 1.001–1.003]), potassium (OR 2.02 [95% CI 1.33–3.08]) and fasting plasma glucose (OR 1.04 [95% CI 1.02–1.07]) significantly improved the discriminatory power of the model to AUC 0.86 (95% CI 0.83–0.90).
Conclusions: Among patients with type 2 diabetes, a combination of past medical and drug history and pulse oximetry data, with four non-expensive laboratory measures, was significantly associated with in-hospital COVID-19 mortality.

Introduction

Diabetes is one of the most frequent comorbidities in patients who are hospitalised for coronavirus disease 2019 (COVID-19).[1] Previous systematic reviews have demonstrated that diabetes is a risk factor for severe disease and is associated with an approximately 2–3 fold increased mortality rate from COVID-19 compared

* Corresponding author. Prevention of Metabolic Disorders Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, No. 24, Yemen Street, Shahid Chamran Highway, P.O. Box: 19395-4763, Tehran, Iran. Tel.: +982122432500; fax: +982122416264.
E-mail address: fzhadaegh@endocrine.ac.ir (F. Hadaegh).

https://doi.org/10.1016/j.puhe.2021.11.007
0033-3506/

with patients without diabetes.[2–5] Results of studies among patients with diabetes have shown that some phenotypic characteristics, radiological and laboratory parameters have been associated with the severity of COVID-19;[6,7] however, diabetes is a heterogeneous disease, and specific phenotypes associated with poorer outcome are inconsistent among studies. In addition, model development studies that predict outcomes among patients with diabetes are sparse due to insufficient sample sizes.[8,9]

In Iran, more than 4,580,000 confirmed cases of COVID-19 and 100,255 deaths had been reported (until 20 August 2021), according to the World Health Organisation (WHO) report.[10] Furthermore, compared with other countries in the Middle East and North Africa (MENA) region, Iran has the highest total number of COVID-19 deaths (as of 20 August 2021).[11] A multicentre, cross-sectional study conducted in 19 hospitals in Tehran, Iran, showed a case fatality rate (CFR) of 10.05% among 16,000 cases of COVID-19.[12] In that study, the highest rate of mortality was observed in patients with diabetes. In another single-centre study including 2968 Iranian patients who were hospitalised with COVID-19, patients with diabetes had significantly higher rates of CFR compared with patients who had no comorbidities (9.73% vs 7.61%).[13]

Globally, in 2017, the MENA region had the second-highest prevalence of type 2 diabetes (10.8%), with an increasing trend of 1.5–2 times in the past three decades.[14] Hence, it was expected that during the COVID-19 pandemic, patients with diabetes in this region would be greatly impacted.

As the burden of disease due to diabetes[15] and COVID-19[10] increases in Iran, the current study aims to: (1) describe the clinical and laboratory characteristics of patients with diabetes and COVID-19; (2) identify the risk factors of in-hospital mortality among these patients; and (3) develop a predictive model for in-hospital mortality among Iranian adult patients with type 2 diabetes who were hospitalised for COVID-19.

Methods

Study population and data collection

The study population included all adult patients (aged ≥18 years) with type 2 diabetes ($n = 560$) who were hospitalised for COVID-19, according to the algorithms suggested by the WHO,[16] at a tertiary referral centre in Golestan province, Iran, between February and August 2020.

Two physicians extracted demographic data, medical and drug history, symptoms and signs, and laboratory parameters from electronic medical records. All inpatient medical records were then completed by telephone calls. Unfortunately, data were not complete for all patients. Details of missing data for each characteristic are shown in Table 1.

Clinical and laboratory measurements

Oropharyngeal swab specimens were collected and examined in predetermined laboratories across the province to detect SARS-CoV-2 viral nucleic acid using a real-time reverse transcription-polymerase chain reaction (RT-PCR) assay. Laboratory parameters, including white blood cells (WBC) count, neutrophils and lymphocytes, haemoglobin (Hb), blood urea nitrogen (BUN) concentration, creatinine, sodium, potassium, creatine phosphokinase (CPK), lactate dehydrogenase (LDH), albumin, liver enzymes (including aspartate and alanine transaminases [AST and ALT]), erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) were collected for each patient. The primary outcome was in-hospital COVID-19 mortality.

Definition of terms

As suggested by the National Headquarters for the management and control of the novel coronavirus, we followed the WHO interim guidelines for diagnosing COVID-19 infection;[17] thus, the case definition was based on both confirmed (i.e. positive PCR) and probable infected cases. Probable cases were defined as follows: first, either a febrile acute respiratory illness (ARI) with clinical, radiological or histopathological evidence of pulmonary parenchymal disease (e.g. pneumonia or acute respiratory distress syndrome [ARDS]), a direct epidemiological link to a laboratory-confirmed COVID-19 case, or testing for COVID-19 is unavailable or negative on a single inadequate specimen or shows inconclusive results; second, a febrile ARI that is not explained fully by any other aetiology, and the person resides or travelled in the Middle East, and testing for COVID-19 is inconclusive; third, an ARI of any severity, and a direct epidemiological link to a confirmed COVID-19 case, and testing for COVID-19 is inconclusive. Diabetes was defined by fasting plasma glucose (FPG) ≥7.0 mmol/L or random blood glucose ≥11.0 mmol/L or glycosylated haemoglobin A1c (HbA1c) ≥6.5% at admission or the patient is already receiving glucose-lowering medications.

Body mass index (BMI) was calculated as weight in kilograms divided by the square of height (m^2). The estimated glomerular filtration rate (eGFR) was expressed in mL/min/1.73 m^2 and was calculated using the chronic kidney diseases (CKD) Epidemiology Collaboration (CKD-EPI) equation,[18] and CKD was defined as an eGFR of <60 mL/min per 1.73 m^2.[19] The education status was classified as illiterate/elementary school, below diploma/diploma and higher than a diploma.

Statistical analyses

The baseline characteristics were presented as mean (standard deviation [SD]) or median (interquartile range [IQR]) for continuous variables and frequencies (%) for categorical variables. Comparisons of baseline characteristics between patients who survived and those who died and between respondents (those with complete data on covariates) and non-respondents (those with missing data on some covariates at the baseline) were performed using Student's *t*-test, Mann–Whitney test and Chi-squared test, as appropriate.

Using univariable logistic regression, the associations of different characteristics of patients at admission with in-hospital mortality were evaluated. Covariates with *P*-values <0.20 in univariable analysis were then entered in the multivariable models. Model 1 included demographic data, diabetes-related complications and drug history; model 2 included all significant variables in model 1, plus vital signs and pulse oximetry data; model 3 included significant variables in model 2, plus laboratory test data. Since the history of chronic insulin therapy (with or without oral antidiabetic drugs [OADs]) was an important indicator for the long duration of diabetes and a predictor of mortality among COVID-19 patients, this confounder was included in both models 2 and 3.

The odds ratios (ORs) and 95% confidence intervals (95% CIs) were reported for each predictor in different models.

The area under the receiver-operating characteristic curve (AUC) was used to assess the discrimination of models. According to the Hosmer et al. criteria, the AUC values were categorised as poor (≥0.5 to <0.7), acceptable (≥0.70 to <80), excellent (≥0.80 to <0.90) and outstanding (≥0.90) discriminations.[20] A comparison of AUC values of different models with the same number of participants in the three models (i.e. $n = 456$) was performed using a non-parametric approach proposed by DeLong et al.[21]

In addition, to address overfitting, which mainly occurs in the model building process, the optimism-corrected AUC was

Table 1
Characteristics of diabetic patients hospitalised for COVID-19 infection ($N = 560$).

Characteristics	Patients with available data	Total population	Patient outcome		*P*-value
			Deceased ($n = 165$)	Survived ($n = 395$)	
Demographic characteristics					
Gender (male), *n* (%)	560	291 (52%)	93 (56.4%)	198 (50.1%)	0.19
Age (years), mean ± SD	560	61.8 (13.4)	64.3 (14.0)	60.7 (13.1)	0.004
BMI (kg/m^2), mean ± SD	479	29.2 (6.1)	29.0 (6.5)	29.3 (5.9)	0.50
Marital status (married) *n* (%)	560	551 (98.4%)	162 (98.2%)	389 (98.5%)	0.65
Education, *n* (%)	512				0.76
Illiterate/elementary		317 (61.9%)	98 (62.8%)	219 (61.5%)	
Below diploma/diploma		124 (24.2%)	39 (25%)	85 (23.9%)	
Higher than diploma		71 (13.9%)	19 (12.2%)	52 (14.6%)	
Area of residence, *n* (%)	560				0.22
Rural		157 (28.0%)	40 (24.2%)	117 (29.6%)	
Urban		403 (72.0%)	125 (75.8%)	278 (70.4%)	
Duration of diabetes, mean ± SD	469	8.01 (8.0)	7.4 (8.9)	8.3 (8.5)	0.33
Comorbidities, *n* (%)					
History of hypertension	559	441 (78.9%)	122 (73.9%)	319 (81.0%)	0.07
History of previous CAD	515	220 (42.7%)	63 (42.0%)	157 (43.0%)	0.84
History of stroke	510	53 (10.4%)	23 (15.4%)	30 (8.3%)	0.02
History of pulmonary disease	506	94 (18.6%)	35 (23.6%)	59 (16.5%)	0.08
History of diabetic foot	485	76 (15.7%)	50 (14.9%)	26 (17.4%)	0.50
Routine treatment before admission, *n* (%)					
Antidiabetic drugs	560				0.001
OADs		215 (38.4%)	45 (27.3%)	170 (43%)	
Insulin		101 (18%)	28 (17%)	73 (18.5%)	
Both OADs and insulin		48 (8.6%)	17 (10.3%)	31 (7.8%)	
Beta-blocker	560	76 (13.6%)	15 (9.1%)	61 (15.4%)	0.06
ARBs/ACE inhibitors	560	249 (44.5%)	66 (40.0%)	183 (46.3%)	0.20
Statins	560	111 (19.8%)	21 (12.7%)	90 (22.8%)	0.007
Antiplatelet drugs	560	124 (22.1%)	31 (18.8%)	93 (23.5%)	0.26
Chest CT imaging, *n* (%)	423				0.31
Without involvement		105 (24.8%)	29 (24.4%)	76 (25.0%)	
Crazy paving + consolidation		148 (35.0%)	48 (40.3%)	100 (32.9%)	
Other		170 (40.2%)	42 (35.3%)	128 (42.1%)	
Clinical characteristics, *n* (%)					
Dyspnoea	548	338 (70.8%)	108 (66.3%)	280 (72.7%)	0.15
Cough	538	278 (51.7%)	77 (48.4%)	201 (53.0%)	0.34
Fever	540	275 (50.9%)	90 (56.6%)	185 (48.6%)	9.09
Fatigue	542	255 (47.0%)	84 (52.5%)	171 (44.8%)	0.10
Gastrointestinal symptoms	539	207 (38.4%)	73 (45.9%)	134 (35.3%)	0.02
Cognitive impairment	541	83 (15.3%)	41 (25.8%)	42 (11.0%)	<0.001
Anosmia/hyposmia/ageusia	534	76 (14.2%)	17 (10.8%)	59 (15.7%)	0.17
Vital signs on admission					
SBP (mmHg), mean ± SD	560	137.8 (25.7)	132.2 (26.2)	140.2 (25.2)	0.001
DBP (mmHg), mean ± SD	559	81.8 (14.8)	78.8 (15.4)	83.1 (14.3)	0.002
Pulse (beats/min), mean ± SD	558	99.06 (18.4)	100.7 (19.1)	98.4 (18.1)	0.19
RR-breaths (per minute), median (IQR)	559	20.0 (18.0–25)	24.0 (20–28)	20.0 (18–24)	<0.001
Temperature (°C), median (IQR)	557	37.0 (37.0–37.5)	37.2 (37–37.6)	37.0 (37–37.5)	0.15
SPO2 (%), median (IQR)	560	91.0 (85–95)	85.0 (72.5–91)	93.0 (88–96)	<0.001
Admission plasma glucose (mg/dl), mean ± SD	559	231.4 (114.6)	231.8 (29.0)	231.2 (108.1)	0.95
White blood cell count (× 10^9/L), mean ± SD	560	8.86 (4.8)	9.82 (4.5)	8.46 (4.8)	0.002
Neutrophil count (%), median (IQR)	499	79.0 (70–86)	83.0 (76.5–88.0)	77.0 (69.7–84)	<0.001
Lymphocyte count (%), median (IQR)	499	16.0 (10–24)	12.0 (8–19)	18.0 (12–26)	<0.001
Haemoglobin (g/L), mean ± SD	560	119.0 (19)	118.0 (18.8)	119.0 (19)	0.34
Platelet count (× 10^9/L), mean ± SD	560	214.0 (94.8)	199.5 (90.0)	220.1 (96.1)	0.02
Prothrombin time (s), median (IQR)	283	13.2 (13–14.7)	13.7 (13–15.2)	13.0 (13–14.0)	0.77
Partial thromboplastin time (s), median (IQR)	283	32.0 (27–39)	32.0 (27.2–41)	31.8 (27–37)	0.82
HbA1c (%), median (IQR)	70	9.4 (6.9–10.3)	9.5 (6.4–10.3)	9.4 (7–10.4)	0.68
CRP (mg/L), mean ± SD	496	1.2 (0.4)	1.16 (0.37)	1.18 (0.38)	0.62
ESR (mm/h), median (IQR)	330	66.0 (41–93)	65.0 (41.5–94)	66.0 (41–92.5)	0.76
Albumin (g/dl), mean ± SD	343	3.59 (0.5)	3.39 (0.5)	3.7 (0.5)	<0.001
LDH (U/L), median (IQR)	528	563.5 (426.2–777.0)	767.0 (537–995)	512.0 (406–672)	<0.001
CPK, median (IQR)	528	150.0 (88–297.0)	201.5 (113.7–553.5)	131.0 (81–222)	<0.001
Urea (mmol/L), median (IQR)	559	1.43 (1.03–2.45)	1.82 (1.25–3.14)	1.32 (0.92–2.07)	<0.001
Creatinine (μmol/L), median (IQR)	559	106.1 (88.4–141.4)	114.9 (97.2–176.8)	97.2 (88.4–132.6)	0.001
eGFR (mL/min/1.73 m^2), mean ± SD	559	57.47 (26)	51.0 (27.8)	59.3 (24.8)	0.001
Sodium (mEq/L), mean ± SD	558	136.3 (5.7)	136.2 (7.0)	136.37 (5.0)	0.73
Potassium (mEq/L), mean ± SD	558	4.48 (0.7)	4.6 (0.9)	4.4 (0.7)	0.01
AST (U/L), median (IQR)	305	38.0 (27–65.5)	48.0 (30–91)	54.2 (23.7–58.2)	0.001
ALT (U/L), median (IQR)	305	34.0 (21–56.5)	37.0 (22–73)	30.5 (20–56)	0.014
FPG (mg/dl), mean ± SD	560	250.5 (114.7)	282.2 (133.3)	237.3 (103.3)	<0.001

Abbreviations: SD, standard deviation; IQR, interquartile range; BMI, body mass index; OADs, oral antidiabetic drugs; ARB, angiotensin receptor blocker; ACE, angiotensin-converting enzyme; CAD, coronary artery disease; SBP, systolic blood pressure; DBP, diastolic blood pressure; RR, respiratory rate; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; LDH, lactate dehydrogenase; CPK, creatine phosphokinase; eGFR, estimated glomerular filtration rate (as calculated using the CKD Epidemiology Collaboration equation); AST, aspartate aminotransferase; ALT, alanine aminotransferase; FPG, fasting plasma glucose.

estimated using 1000 bootstrap resamples for every underlying model. The difference between the original and the mean AUC of the 1000 replicates was used as a correction factor and subtracted from the original AUC. This bias-corrected AUC was used as a measure for internal validation.

To evaluate the calibration, which shows agreement between the observed (actual) outcomes and predictions, we used observed to predicted ratios, the Hosmer–Lemeshow goodness-of-fit test and a calibration plot. The calibration plot shows predicted in-hospital death probabilities (x-axis) against the observed outcomes (y-axis) in deciles of the predicted probabilities. Using the LOWESS (locally weighted scatter plot smoothing) line, we smoothed the calibration plot. Perfect predictions are on the 45° line (y = x). Validation of the goodness-of-fit of each underlying model was determined by the Hosmer–Lemeshow test in deciles based on the predicted risk. A non-significant test implied that the observed outcome did not differ significantly from the predicted mortality risk.

To encourage the integration of the prognostic model into everyday clinical situations, the mathematical formula of the prognostic algorithm obtained from logistic regression modelling was also incorporated into a nomogram. The nomogram developed herein serves as a graphical representation of our prognostic algorithm, incorporating significant prognostic factors as continuous variables to predict the risk of in-hospital mortality from COVID-19. Except for the variable selection, $P < 0.05$ was considered significant. Statistical analyses were performed with SPSS 22 (SPSS Inc., Chicago, IL, USA) and STATA 14 (StataCorp, college station, TX, USA).

Results

Comparison between respondents and non-respondents indicated no clinically important differences between these two groups, with the exception that respondents reported a higher frequency of angiotensin-converting enzyme (ACE)/angiotensin II receptor blockers (ARB) use and a lower prevalence of cough and gastrointestinal symptoms compared with non-respondents (see Supplementary Table S1).

The study sample included 560 patients with diabetes (men = 291). Among them, 364 (65%) were receiving glucose-lowering medication. The mean age of the total population was 61.8 (SD 13.4) years, and SARS-CoV-2 PCR testing was performed in 209 patients, with a positive result in 125 patients. The median duration of hospital stay was 6 (IQR 3–11) days, whereby 232 patients were admitted to the intensive care unit (ICU) for the median stay of 6 (IQR 3–11) days. In total, 165 in-hospital deaths were recorded (men = 93).

The baseline characteristics of patients who survived and those who died are compared in Table 1. The prevalence of overweight, obesity and CKD was 37.2%, 38.6% and 44.5%, respectively. A medical history of hypertension, coronary artery disease (CAD), stroke and pulmonary disease was observed in 78.9%, 42.7%, 10.4% and 18.6% of the participants, respectively. The mean level of plasma glucose at the time of hospital admission was 231.4 (114.6) mg/dl, and the level of HbA1c (only for 70 patients) was 9.4%. The most common glucose-lowering medications were metformin, followed by insulin, sulfonylurea and other oral glucose-lowering agents. Moreover, ACE inhibitors and/or ARBs, beta-blockers, statins and antiplatelet drugs were used by 44.5%, 13.6%, 19.8% and 22.1% of the participants, respectively.

The most common signs of COVID-19 on admission were dyspnoea, cough, fever, fatigue, gastrointestinal disorders, cognitive impairment and anosmia, hyposmia and ageusia. Thoracic computed tomography (CT) imaging was performed for all patients at hospital admission and did not reveal any abnormality in 25% of patients. Details of other results are shown in Table 1.

Patients who died compared with those who survived were older, more likely to have a history of stroke, and present with gastrointestinal symptoms and cognitive impairment. Moreover, they were less likely to be taking metformin and statins. In-hospital mortality was more likely in individuals who initially presented (i.e. at hospital admission) with significantly lower systolic blood pressure (SBP), diastolic blood pressure (DBP), oxygen saturation (SpO2) and lower levels of lymphocytes, platelet, albumin and eGFR, but higher levels of respiratory rate, WBC, neutrophils, LDH, CPK, creatinine, potassium, AST, ALT and FPG compared with patients who survived (all *P*-values were <0.05).

Table 2 shows multivariate prediction models for in-hospital mortality. In model 1, age (OR 1.02 [95% CI 1.00–1.04]) and a history of cognitive impairment (OR 3.17 [95% CI 1.77–5.68]) were associated with a significantly higher risk of in-hospital mortality. Moreover, prior use of OADs, beta-blockers and statins was associated with significant 55%, 51% and 49% lower risks of mortality, respectively. In model 2, age and history of cognitive impairment were independently associated with a higher risk of mortality, while the use of statins, beta-blockers and OADs, lower respiratory rate (OR 1.07 [95% CI 1.00–1.14]) and higher oxygen saturation (OR 0.95 [95% CI 0.92–0.98]) were associated with a significantly lower risk of mortality. Finally, in model 3, in addition to the significant predictors of model 2, the use of insulin (OR 0.42 [95% CI 0.19–0.94]), platelet count (OR 0.99 [95% CI 0.99–1.00]), LDH (OR 1.002 [95% CI 1.001–1.003]), potassium (OR 2.02 [95% CI 1.33–3.08]) and each 10 mg/dl increase in FPG (OR 1.04 [95% CI 1.04–1.07]) were found to be independently associated with the risk of death.

The values of discrimination power (AUC) of models were 0.75 (95% CI 0.70–0.80) for model 1, 0.80 (95% CI 0.74–0.82) for model 2, and 0.86 (95% CI 0.83–0.90) for model 3. The corrected AUC for model 3 was 0.82 (95% CI 0.79–0.89). Model 3 had the highest discrimination power (Fig. 1).

The calibration plot indicated good calibration for the risk prediction model within the data set (Fig. 2) and the Hosmer–Lemeshow goodness-of-fit test also showed good calibration ($\widehat{C} = 4.19$, $P = 0.84$).

Fig. 3 shows the nomogram of the final model (model 3). According to the nomogram as an example, a 45-year-old patient with diabetes who presents at the emergency room with an SpO2 of 80%, FPG of 210 mg/dl, LDH of 800 U/L, potassium of 5 mEq/L and platelet count of 230×10^9/L, without a history of loss of consciousness and who is not receiving statin, beta-blocker or OAD treatment gets the score of ($1.5 + 1 + 1 + 2 + 3.5 + 4 + 3 + 0 + 1 + 1.2 + 1 = 19.2$) and will have a 95% probability of mortality.

Discussion

The current study was conducted in a large tertiary centre in the North East of Iran during the first half of 2020. Our findings, among 560 patients with diabetes who were hospitalised for COVID-19, showed a 30% in-hospital mortality rate following approximately one week of hospitalisation. Ageing, cognitive impairment and higher levels of LDH, potassium and FPG were found to be associated with an increased risk of death, while higher platelet levels and oxygen saturation, as well as taking oral glucose-lowering drugs, insulin, statins and beta-blockers, were significantly associated with a reduced risk of in-hospital mortality. We developed a simple model with its nomogram that showed an excellent discriminatory power for the prediction of mortality events among patients with diabetes who had been hospitalised with COVID-19.

Table 2
Multivariate prediction models of in-hospital mortality for patients with diabetes and COVID-19.

Characteristics	Model 1 ($N = 498$)		Model 2 ($N = 534$)		Model 3 ($N = 456$)	
	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
Prior to admission characteristics						
Age	1.02 (1.00–1.04)	0.01	1.02 (1.00–1.04)	0.01	1.03 (1.01–1.05)	0.01
Gender	1.01 (0.65–1.56)	0.97				
Area of residence	0.74 (0.45–1.19)	0.19				
Dyspnoea	0.66 (0.42–1.04)	0.07				
History of fever	1.53 (0.96–2.43)	0.07				
Cognitive impairment	3.17 (1.77–5.68)	<0.001	2.06 (1.16–3.66)	0.01	2.78 (1.35–5.71)	0.006
Fatigue	1.34 (0.85–2.13)	0.21				
Gastrointestinal symptoms	1.56 (0.97–2.51)	0.06				
Anosmia, hyposmia or ageusia	0.54 (0.28–1.04)	0.06				
History of stroke	1.22 (0.60–2.49)	0.58				
History of pulmonary disease	1.51 (0.88–2.58)	0.13				
History of hypertension	0.63 (0.36–1.12)	0.12				
Routine treatment before admission						
OADs	0.45 (0.26–0.75)	0.003	0.55 (0.33–0.92)	0.02	0.38 (0.12–0.73)	0.04
Insulin	0.53 (0.28–1.01)	0.055	0.83 (0.45–1.52)	0.54	0.42 (0.19–0.94)	0.03
OADs and insulin	0.97 (0.45–2.09)	0.94	1.24 (0.57–2.71)	0.58	1.01 (0.38–2.70)	0.99
Beta-blocker	0.49 (0.24–0.99)	0.05	0.43 (0.20–0.89)	0.02	0.35 (0.14–0.88)	0.02
ARBs/ACE inhibitors	0.98 (0.57–1.69)	0.95				
Statins	0.51 (0.28–0.93)	0.03	0.46 (0.24–0.86)	0.01	0.37 (0.17–0.82)	0.01
Vital signs on admission						
SBP			0.99 (0.97–1.00)	0.06		
DBP			1.007 (0.98–1.03)	0.52		
Pulse			0.99 (0.98–1.00)	0.16		
RR			1.07 (1.00–1.14)	0.04	1.06 (0.97–1.15)	0.18
Temperature			1.02 (0.80–1.45)	0.62		
SPO2			0.95 (0.92–0.98)	0.001	0.95 (0.91–0.99)	0.01
Laboratory tests on admission						
White blood cell count					0.98 (0.93–1.03)	0.48
Neutrophil count					0.99 (0.90–1.08)	0.80
Lymphocyte count					0.97 (0.87–1.07)	0.54
Platelet					0.99 (0.99–1.00)	0.002
LDH					1.002 (1.001–1.003)	0.00
Creatinine					1.07 (0.78–1.47)	0.68
eGFR					1.00 (0.98–1.01)	0.64
Potassium					2.02 (1.33–3.08)	0.001
FPG					1.04 (1.02–1.07)	0.001
AUC	0.76 (0.70–0.80)		0.80 (0.74–0.82)		0.86 (0.83–0.90)	

Abbreviations: OADs, oral antidiabetic drugs; ARB, angiotensin receptor blocker; ACE, angiotensin-converting enzyme; SBP, systolic blood pressure; DBP, diastolic blood pressure; RR, respiratory rate; LDH, lactate dehydrogenase; eGFR, estimated glomerular filtration rate (as calculated using the CKD Epidemiology Collaboration equation); FPG, fasting plasma glucose; AUC, area under the receiver-operating characteristic curve; OR, odds ratio; CI, confidence interval.
The ORs correspond to each one-unit increase for continuous variables except for 10 mg/dL increase in FPG.

Several studies[6,8,22–25] have assessed the characteristics and prognostic factors among the diabetic population with different results (Supplementary Table S2). Two studies specifically focused on patients with diabetes who were hospitalised for COVID-19. First, the CORONADO study[7] showed a 10.6% risk of death and found several factors, including age, treated obstructive sleep apnoea and microvascular and macrovascular complications, to be independent predictors of death on day 7 (the current study observed a higher risk of mortality, and the only common predictor of death with our population was age). The other study among patients with diabetes who were hospitalised with COVID-19 was in the US[6] and showed a mortality rate of 33.1%, which is comparable with the results obtained in our study. Furthermore, the US study showed that HbA1c was not associated with mortality events, while insulin treatment was a strong predictor of mortality. In the current study, we found that a high level of plasma glucose at hospital admission, as a proxy for the level of diabetes control,[26] was significantly associated with an increased risk of mortality. Moreover, in contrast to the study of Agarwal et al., we demonstrated that a history of using OADs and insulin was significantly associated with a lower risk of in-hospital mortality. Importantly, according to a national study,[27] despite some improvement in the knowledge and screening of diabetes in Iran, 24.7% of patients with diabetes were not aware of their disease. In addition, these newly diagnosed patients exhibited a coronary heart disease (CHD) risk comparable to patients without diabetes with a prior CHD event.[28] Results of a systematic review and meta-analysis of 14 studies showed that male sex, age, history of cardiovascular disease (CVD), CKD, chronic obstructive pulmonary disease (COPD), high plasma glucose at hospital admission and chronic insulin use were associated with a high risk of death for patients with diabetes who also had COVID-19.[25]

Among patients with diabetes, a few studies found no association between statin use and poor outcome.[7,22,25,29] Several studies[30,31] with larger sample sizes among patients without diabetes showed that previous statin use in patients hospitalised with COVID-19 was associated with lower in-hospital mortality, which might be related to the immunomodulatory action or by preventing cardiovascular damage in addition to their lipid-lowering activity.[32] These results are consistent with our study that also found a 63% lower risk of death among patients who used statins, despite the low number of lipid-lowering medication users in our study population. In another study, low statin use was also observed among patients with diabetes in Iran.[33] It is interesting that the current study found that using beta-blockers was significantly associated with a lower risk of in-hospital mortality. The results of other

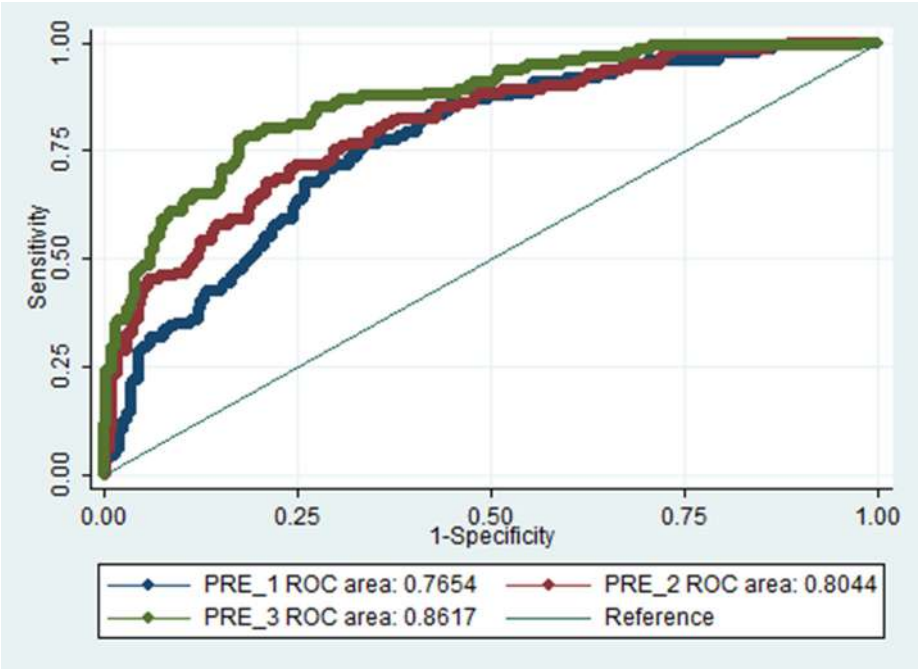
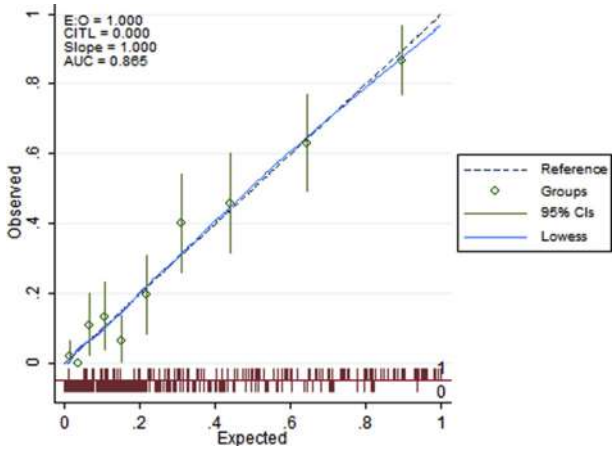


Fig. 1. Receiver-operating characteristic curve for the three models. Blue curve, model 1 (age, history of cognitive impairment, use of statins, beta-blockers and oral antidiabetic drugs); red curve, model 2 (age, history of cognitive impairment, use of statins, beta-blockers and oral antidiabetic drugs, respiratory rate and oxygen saturation); and green curve, model 3 (age, history of cognitive impairment, use of statins, beta-blockers, oral antidiabetic drugs, insulin, oxygen saturation, platelet count, lactate dehydrogenase, potassium and fasting plasma glucose). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

studies among patients with diabetes who were infected with COVID-19 were inconsistent; one study showed a 19% higher risk, and another study found a 33% lower risk of mortality risk among beta-blocker users, although both associations were non-significant.[7,22] According to different guidelines, using this group of medications for patients with diabetes is limited to those with acute coronary syndrome or who are experiencing heart failure (HF).[34] Systematic reviews revealed that using beta-blockers was



Deciles	1	2	3	4	5	6	7	8	9	10
Observed risk	0.022	0.00	0.109	0.133	0.065	0.195	0.400	0.456	0.630	0.866
Expected risk	0.014	0.035	0.064	0.106	0.152	0.216	0.309	0.441	0.644	0.893
Ratio	1.55	0.0	1.69	1.26	0.43	0.90	1.29	1.03	0.98	0.97

Fig. 2. Calibration curves for comparing actual and predicted in-hospital death probabilities according to model 3. The Hosmer–Lemeshow test for goodness-of-fit ($\hat{C} = 4.19$, $P = 0.84$).

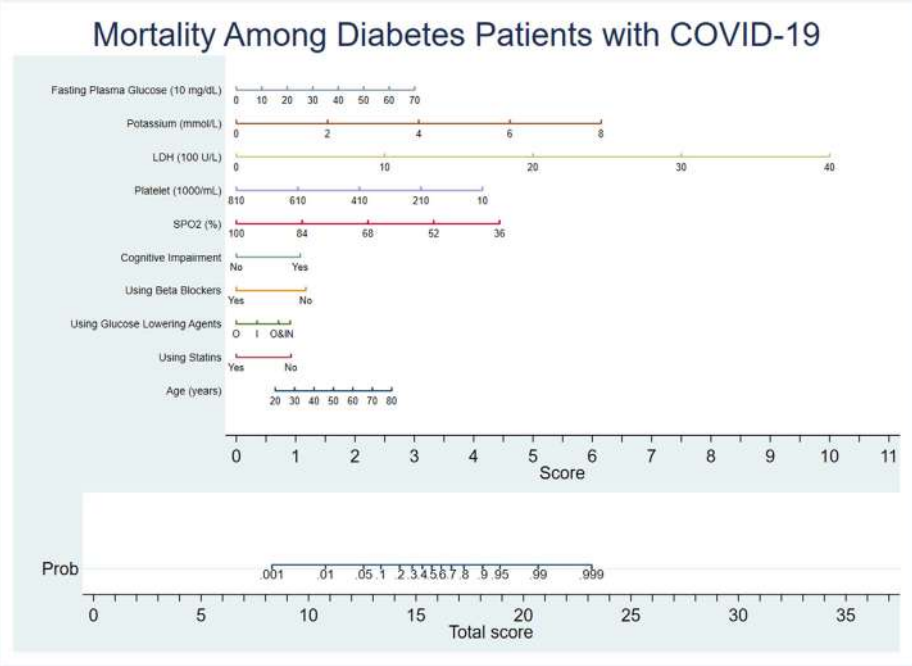


Fig. 3. Nomogram of covariates in model 3 for predicting in-hospital mortality among patients with diabetes who are hospitalised with COVID-19. Upon hospital admission, the risk of mortality for a patient with diabetes can be calculated by computing the corresponding score of points for each of the nine clinical characteristics and then adding them together. Looking at the bottom scale, the probability of in-hospital mortality corresponding to the calculated score can be calculated. LDH, lactate dehydrogenase; Prob, probability; O, oral antidiabetic drugs; I, insulin; O&I, oral antidiabetic drugs and insulin; N, no antidiabetic drugs.

associated with improved outcomes among patients with HF, regardless of the diabetes status.[35] Hence, in the current study population, with a prevalence of CVD >40%, we speculated that a large number of patients had some degree of existing HF and would benefit from beta-blockers. Unfortunately, echocardiography was not performed for most patients; hence, the ejection fraction data were not available.

To the best of our knowledge, no study has investigated the association between potassium level and COVID-19 mortality among patients with diabetes; however, in studies conducted in the general population, potassium was not associated with COVID-19 mortality risk[36] or a higher risk[37] of COVID-19 infection. In the current study, each 1 mEq/L higher potassium level was associated with a two-fold higher risk of in-hospital mortality. These findings are in line with previous studies among ICU patients showing that hyperkalaemia is an independent risk factor for death, even at a moderate increase above the normal levels.[38,39] Importantly, the significant risk of a higher potassium level, which was found in the current study, was independent of several important confounding factors, including eGFR and ACE inhibitor/ARB use. However, pH value is a strong confounder for potassium level as metabolic acidosis can cause a potassium shift from the intracellular to the extracellular space; unfortunately, we did not have data on pH levels. The most important mechanism that is caused by hyperkalaemia is lowering of the resting membrane potential of the myocardium, resulting in decreased myocardial cell conduction velocity and increased rate of repolarisation.[39]

To date, only two studies conducted among Chinese and Australian populations have presented prediction models for mortality from COVID-19 among patients with diabetes.[8,9] The study among the Chinese population[9] developed a model containing predictor variables, including partial thromboplastin time (PTT), urea nitrogen, WBC count and LDH, with an AUC of 0.836. Moreover, the study in Australia[8] showed that a score computed from age, arterial occlusive disease, eGFR, CRP and AST levels at hospital admission predicted in-hospital mortality with a C statistic of 0.89 and good calibration. Our prediction model with 11 variables, including four non-expensive laboratory measures, yielded similar excellent discriminatory power (0.86 [95% CI 0.83–0.90]) with acceptable calibration.

Strengths of the current study were the large sample size of patients who were hospitalised for COVID-19 and a considerable number of patients with diabetes, which, as discussed previously, imposes a great burden of disease in Iran.[10,15]

The current findings need to be interpreted in light of the study limitations. First, this study did not validate the model externally; however, the model showed a reasonable internal validity as examined by the bootstrapping method. Second, there were no data on HbA1c; however, we adjusted for plasma glucose as a surrogate of the HbA1c level.[26] Third, a large number of patients were probable cases of COVID-19, and the cases were not confirmed with PCR. However, previous studies have shown that because of difficulty in obtaining reliable nasopharyngeal swab specimens, the timing of detection and limited detection capacity during the early stages of the outbreak, false-negative results are often seen in the PCR method.[40] Moreover, CT imaging of the chest is a more reliable, feasible and rapid method to diagnose and assess COVID-19 compared with RT-PCR, especially in epidemic regions such as Iran.[41–43] Fourth, we did not categorise patients as those with previous diabetes and those who had increased blood glucose due to COVID-19 infection, which might overestimate the number of patients with diabetes. Finally, data regarding in-hospital treatment were not included in our analysis.

In conclusion, approximately one-third of patients with diabetes who were hospitalised for COVID-19 in this large referral centre located in the North East of Iran died within 1 week of admission. A simple and non-expensive risk score consisting of 11 variables, including age, history of cognitive impairment, use of statins, OADs,

insulin, beta-blockers, SpO2, platelet count, LDH, potassium and FPG levels, demonstrated excellent prediction for in-hospital mortality among patients with diabetes. This simple risk score may help physicians in emergency departments to assess the prognosis of patients with diabetes.

Author statements

Acknowledgements

The authors appreciate the efforts of healthcare workers and support staff at Sayyad Shirazi Hospital for providing patient care at considerable personal risk on the front lines of this pandemic. We would also like to express our solidarity with those who are or have been ill with COVID-19 and their families.

Ethics approval

This study was approved by the Ethics Committee of the Golestan University of Medical Science (Protocol no. IR.GOUMS.REC.1399.043) and was in accordance with the 1964 Declaration of Helsinki. We obtained verbal consent for publication from living patients and the relatives of deceased patients.

Funding

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

Competing interests

None declared.

Availability of data and materials

All data and materials are available upon request.

Author contributions

M.K, F.H and R.HT planned the study, researched the data and wrote the manuscript. M.H and MR.B analysed the data, reviewed and edited the manuscript. D.K, H.A and GR.R reviewed and edited the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2021.11.007.

References

1. Atkins JL, Masoli JA, Delgado J, Pilling LC, Kuo C-LC, Kuchel G, et al. Preexisting comorbidities predicting severe COVID-19 in older adults in the UK biobank community cohort. *medRxiv* 2020;**75**:2224–30.
2. Huang I, Lim MA, Pranata R. Diabetes mellitus is associated with increased mortality and severity of disease in COVID-19 pneumonia–a systematic review, meta-analysis, and meta-regression. *Diabetes Metab Syndr: Clin Res Rev* 2020;**14**:395–403.
3. Mantovani A, Byrne CD, Zheng M-H, Targher G. Diabetes as a risk factor for greater COVID-19 severity and in-hospital death: a meta-analysis of observational studies. *Nutr Metabol Cardiovasc Dis* 2020;**30**:1236–48.
4. Kumar A, Arora A, Sharma P, Anikhindi SA, Bansal N, Singla V, et al. Is diabetes mellitus associated with mortality and severity of COVID-19? A meta-analysis. *Diabetes Metab Syndr: Clin Res Rev* 2020;**14**:535–45.
5. Guo L, Shi Z, Zhang Y, Wang C, Moreira NCDV, Zuo H, et al. Comorbid diabetes and the risk of disease severity or death among 8807 COVID-19 patients in China: a meta-analysis. *Diabetes Res Clin Pract* 2020;**166**:108346.
6. Agarwal S, Schechter C, Southern W, Crandall JP, Tomer Y. Preadmission diabetes-specific risk factors for mortality in hospitalized patients with diabetes and coronavirus disease 2019. *Diabetes Care* 2020;**43**:2339–44.
7. Cariou B, Hadjadj S, Wargny M, Pichelin M, Al-Salameh A, Allix I, et al. Phenotypic characteristics and prognosis of inpatients with COVID-19 and diabetes: the CORONADO study. *Diabetologia* 2020:1–16.
8. Sourij H, Aziz F, Bräuer A, Ciardi C, Clodi M, Fasching P, et al. COVID-19 fatality prediction in people with diabetes and prediabetes using a simple score upon hospital admission. *Diabetes Obes Metabol* 2021;**23**:589–98.
9. Su M, Yuan J, Peng J, Wu M, Yang Y, Peng YG. Clinical prediction model for mortality of adult diabetes inpatients with COVID-19 in Wuhan, China: a retrospective pilot study. *J Clin Anesth* 2020;**66**:109927.
10. https://www.who.int/countries/irn/.
11. https://covid19.healthdata.org/iran-(islamic-republic-of)?view=total-deaths&tab=compare.
12. Zali A, Gholamzadeh S, Mohammadi G, Looha MA, Akrami F, Zarean E, et al. Baseline characteristics and associated factors of mortality in COVID-19 Patients; an analysis of 16000 cases in Tehran, Iran. *Arch Acad Emerg Med* 2020;**8**.
13. Nikpouraghdam M, Farahani AJ, Alishiri G, Heydari S, Ebrahimnia M, Samadinia H, et al. Epidemiological characteristics of coronavirus disease 2019 (COVID-19) patients in Iran: a single center study. *J Clin Virol* 2020;**127**:104378.
14. Azizi F, Hadaegh F, Hosseinpanah F, Mirmiran P, Amouzegar A, Abdi H, et al. Metabolic health in the Middle East and north Africa. *Lancet Diabetes Endocrinol* 2019;**7**:866–79.
15. Esteghamati A, Gouya MM, Abbasi M, Delavari A, Alikhani S, Alaedini F, et al. Prevalence of diabetes and impaired fasting glucose in the adult population of Iran: national survey of risk factors for non-communicable diseases of Iran. *Diabetes Care* 2008;**31**:96–8.
16. Organization WH. *Algorithm for COVID-19 triage and referral: patient triage and referral for resource-limited settings during community transmission*. Manila: WHO Regional Office for the Western Pacific; 2020.
17. Organization WH. *Clinical management of severe acute respiratory infection when novel coronavirus (nCoV) infection is suspected: interim guidance, 25 January 2020*. World Health Organization; 2020.
18. Levey AS, Stevens LA, Schmid CH, Zhang Y, Castro III AF, Feldman HI, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009;**150**:604–12.
19. Levey AS, Eckardt K-U, Tsukamoto Y, Levin A, Coresh J, Rossert J, et al. Definition and classification of chronic kidney disease: a position statement from Kidney Disease: improving Global Outcomes (KDIGO). *Kidney Int* 2005;**67**:2089–100.
20. Hosmer Jr DW, Lemeshow S, Sturdivant RX. *Applied logistic regression*. John Wiley & Sons; 2013.
21. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988:837–45.
22. Orioli L, Servais T, Belkhir L, Laterre P-F, Thissen J-P, Vandeleene B, et al. Clinical characteristics and short-term prognosis of in-patients with diabetes and COVID-19: a retrospective study from an academic center in Belgium. *Diabetes Metab Syndr: Clin Res Rev* 2021;**15**:149–57.
23. Chen Y, Yang D, Cheng B, Chen J, Peng A, Yang C, et al. Clinical characteristics and outcomes of patients with diabetes and COVID-19 in association with glucose-lowering medication. *Diabetes Care* 2020;**43**:1399–407.
24. Shi Q, Zhang X, Jiang F, Zhang X, Hu N, Bimu C, et al. Clinical characteristics and risk factors for mortality of COVID-19 patients with diabetes in Wuhan, China: a two-center, retrospective study. *Diabetes Care* 2020;**43**:1382–91.
25. Schlesinger S, Neuenschwander M, Lang A, Pafili K, Kuss O, Herder C, et al. *Risk phenotypes of diabetes and association with COVID-19 severity and death—a living systematic review and meta-analysis*. 2020.
26. Association AD. 6. Glycemic targets. *Diabetes Care* 2015;**38**:S33–40.
27. Esteghamati A, Etemad K, Koohpayehzadeh J, Abbasi M, Meysamie A, Noshad S, et al. Trends in the prevalence of diabetes and impaired fasting glucose in association with obesity in Iran: 2005–2011. *Diabetes Res Clin Pract* 2014;**103**:319–27.
28. Hadaegh F, Fahimfar N, Khalili D, Sheikholeslami F, Azizi F. New and known type 2 diabetes as coronary heart disease equivalent: results from 7.6 year follow up in a Middle East population. *Cardiovasc Diabetol* 2010;**9**:1–8.
29. Fadini GP, Morieri ML, Boscari F, Fioretto P, Maran A, Busetto L, et al. Newly-diagnosed diabetes and admission hyperglycemia predict COVID-19 severity by aggravating respiratory deterioration. *Diabetes Res Clin Pract* 2020;**168**:108374.
30. Gupta A, Madhavan MV, Poterucha TJ, DeFilippis EM, Hennessey JA, Redfors B, et al. Association between antecedent statin use and decreased mortality in hospitalized patients with COVID-19. *Nat Commun* 2021;**12**:1–9.
31. Tan WY, Young BE, Lye DC, Chew DE, Dalan R. Statin use is associated with lower disease severity in COVID-19 infection. *Sci Rep* 2020;**10**:1–7.
32. Castiglione V, Chiriacò M, Emdin M, Taddei S, Vergaro G. Statin therapy in COVID-19 infection. *Eur Heart J Cardiovasc Pharmacother* 2020;**6**:258–9.
33. Jahangiri-Noudeh Y, Akbarpour S, Lotfaliany M, Zafari N, Khalili D, Tohidi M, et al. Trends in cardiovascular disease risk factors in people with and without diabetes mellitus: a Middle Eastern cohort study. *PLos One* 2014;**9**:e112639.
34. Ettehad D, Emdin CA, Kiran A, Anderson SG, Callender T, Emberson J, et al. Blood pressure lowering for prevention of cardiovascular disease and death: a systematic review and meta-analysis. *Lancet* 2016;**387**:957–67.

35. Shibata MC, Flather MD, Wang D. Systematic review of the impact of beta blockers on mortality and hospital admissions in heart failure. *Eur J Heart Fail* 2001;**3**:351–7.
36. Mudatsir M, Fajar JK, Wulandari L, Soegiarto G, Ilmawan M, Purnamasari Y, et al. Predictors of COVID-19 severity: a systematic review and meta-analysis. *F1000Research* 2020;**9**.
37. Garcia PDW, Fumeaux T, Guerci P, Heuberger DM, Montomoli J, Roche-Campo F, et al. Prognostic factors associated with mortality risk and disease progression in 639 critically ill patients with COVID-19 in Europe: initial report of the international RISC-19-ICU prospective observational cohort. *EClinicalMedicine* 2020;**25**:100449.
38. Engelhardt LJ, Balzer F, Müller MC, Grunow JJ, Spies CD, Christopher KB, et al. Association between potassium concentrations, variability and supplementation, and in-hospital mortality in ICU patients: a retrospective analysis. *Ann Intensive Care* 2019;**9**:1–11.
39. McMahon GM, Mendu ML, Gibbons FK, Christopher KB. Association between hyperkalemia at critical care initiation and mortality. *Intensive Care Med* 2012;**38**:1834–42.
40. Li Y, Yao L, Li J, Chen L, Song Y, Cai Z, et al. Stability issues of RT-PCR testing of SARS-CoV-2 for hospitalized patients clinically diagnosed with COVID-19. *J Med Virol* 2020;**92**:903–8.
41. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology* 2020:200642.
42. Davarpanah AH, Mahdavi A, Sabri A, Langroudi TF, Kahkouee S, Haseli S, et al. Novel screening and triage strategy in Iran during deadly COVID-19 epidemic; value of humanitarian teleconsultation service. *J Am Coll Radiol* 2020;**17**:734.
43. Rastad H, Karim H, Ejtahed H-S, Tajbakhsh R, Noorisepehr M, Babaei M, et al. *Risk and predictors of in-hospital mortality from COVID-19 in patients with diabetes and cardiovascular disease*. 2020.

Public Health 202 (2022) 74–75

Contents lists available at ScienceDirect

Public Health

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

Corrigenda

Corrigendum to 'Impact of COVID-19 on birth rate trends in the Italian Metropolitan Cities of Milan, Genoa and Turin' [Public Health 198 (2021) 35–36]

A.F. De Rose, F. Ambrosini*, G. Mantica, C. Terrone

Department of Urology, Policlinico San Martino, Largo Rosanna Benzi, Genoa, Italy

The authors regret that an error was made in the data collection regarding the number of births in the city of Milan from November 2019 to January 2020. The original published paper reported this as 4187 births, but the correct figure is 2656 births. Thus the following corrections are necessary:

Correction 1, abstract, results:

The sentence:

Birth rates in the cities of Milan, Genoa and Turin decreased by 55%, 12%, and 33%, respectively.

Should read:

Birth rates in the cities of Milan, Genoa, and Turin decreased by 12.4%, 12%, and 33%, respectively.

Correction 2, main text:

The passage:

From November 2019 to January 2020, 1579 births were registered in the City of Turin, while, during the same period of the following year, 1043 were recorded; thus, 536 fewer births (33% decline). Similarly, in the City of Milan, 4187 and 2325 births were recorded from November 2019 to January 2020 and during the same quarter of 2020–2021, respectively (55% % reduction).

Should read:

From November 2019 to January 2020, 1579 births were registered in the City of Turin, while, during the same period of the following year, 1043 were recorded; thus, 536 fewer births (33% decline). Similarly, in the City of Milan, 2656 and 2325 births were recorded from November 2019 to January 2020 and during the same quarter of 2020–2021, respectively (12.4% reduction).

DOI of original article: https://doi.org/10.1016/j.puhe.2021.06.026.

* Corresponding author. Department of Urology, Policlinico San Martino Hospital, University of Genoa, Largo Rosanna Benzi, 10, Genoa, 16132 Italy. Tel.: +3480497079.
E-mail address: f.ambrosini1@gmail.com (F. Ambrosini).

https://doi.org/10.1016/j.puhe.2021.11.009

Correction 3, figure 1:

A corrected Fig. 1 follows

The authors would like to apologise for any inconvenience caused.

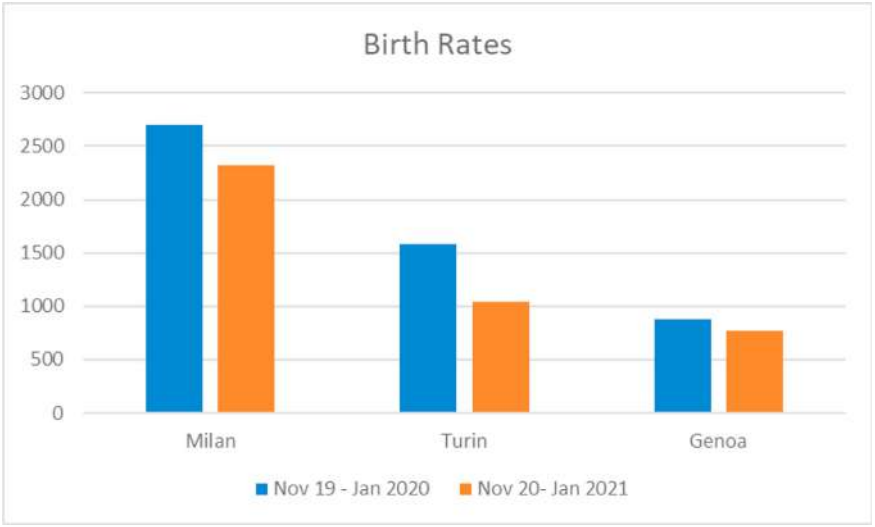


Fig. 1. Birth rates from November 2019 to January 2020 (before the COVID-19 pandemic) and during the same period of the following year (during the COVID-19 pandemic) in the cities of Milan, Genoa, and Turin (Italy).

Public Health 202 (2022) 65

Contents lists available at ScienceDirect

Public Health

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

Corrigenda

Corrigendum to 'Risk perception and resource scarcity in food procurement during the early outbreak of COVID-19' [Public Health 195 (2021) 152–157]

Y. Wang [a], X. Chen [b, c, *], Y. Yang [d], Y. Cui [b], R. Xu [c, e]

[a] Department of Marketing and Management, State University of New York at Oswego, Oswego, NY, 13126, USA
[b] Department of Geography, University of Connecticut, Storrs, CT, 06269, USA
[c] Institute for Collaboration on Health, Intervention, and Policy (InCHIP), USA
[d] Department of Tourism and Hospitality Management, Temple University, Philadelphia, PA, 19122, USA
[e] Department of Allied Health Sciences, University of Connecticut, Storrs, CT, 06269, USA

The authors regret the following errors in the originally published version:

1. In Table 1, for the 'food security status', the 'Secure - secure' category should be '391, 16.37%.'

Here is the updated Table 1:

Table 1
Demographics of survey participants.

Variable	Subgroup	*N* (Percentage)
Gender	Male	1,359 (56.9%)
	Non-male	1,029 (43.1%)
Age	18–24	156 (6.53%)
	25–34	1,039 (43.51%)
	35–44	595 (24.92%)
	45–54	365 (15.28%)
	55–64	177 (7.41%)
	65 and above	56 (2.35%)
Ethnicity	Caucasian	1,692 (70.85%)
	African American	342 (14.32%)
	Latino	129 (5.40%)
	Asian	160 (6.70%)
	Native American	37 (1.55%)
	Other	28 (1.17%)
Educational attainment	Finished middle school	8 (.34%)
	Finished high school	201 (8.42%)
	Some college	398 (16.67%)
	Completed 2-year college	220 (9.21%)
	Completed 4-year college	1,278 (53.52%)
	Attended graduate school	283 (11.85%)
Employment	Employed for wages	1,998 (83.67%)
	Not employed for wages	390 (16.33%)
Food security status (before-during the pandemic)	Secure–secure	391 (16.37%)
	Secure-insecure	438 (18.34%)
	Insecure-secure	24 (1.01%)
	Insecure–insecure	1,535 (64.28%)

2. In Line 148, it is currently written as 'Table 3 reveals the relationships among the behavioral changes: the **increase** in in-store safety perception …'

It should be 'Line 148: Table 3 reveals the relationships among the behavioural changes: the decrease in in-store safety perception …'

Here is the updated paragraph:

Table 3 reveals the relationships among the behavioural changes: the decrease in in-store safety perception was associated with both the decrease in shopping frequency ($\beta = .18$, $P < .01$) and the increase in food expenditure ($\beta = -7.00$, $P < .01$). Also, people's food security status during the pandemic further impacted the relationship between shopping frequency and food expenditure, as shown by the interaction term ($\beta = -22.68$, $P < .01$). This result indicates that the mediation effects on food procurement differ among people in different food security statuses.

The authors would like to apologise for any inconvenience caused.

DOI of original article: https://doi.org/10.1016/j.puhe.2021.04.020.
* Corresponding author. Department of Geography, University of Connecticut, Storrs, CT, 06269, USA.
E-mail address: xiang.chen@uconn.edu (X. Chen).

https://doi.org/10.1016/j.puhe.2021.11.008
0033-3506/Published by Elsevier Ltd on behalf of The Royal Society for Public Health.

Public Health 202 (2022) 93–99

Contents lists available at ScienceDirect

Public Health

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

Original Research

COVID-19 and tobacco cessation: lessons from India

M. Arora [a, b, *], G.P. Nazar [a, b], N. Sharma [a], N. Jain [b], F. Davidson [c], S. Mohan [b, d], D. Mohan [e], M.K. Ali [f], V. Mohan [e], N. Tandon [g], K.M. Venkat Narayan [f], D. Prabhakaran [b, d], L. Bauld [c], K. Srinath Reddy [b]

[a] HRIDAY, New Delhi, India
[b] Public Health Foundation of India, Haryana, New Delhi, India
[c] Usher Institute and SPECTRUM Consortium, University of Edinburgh, Edinburgh, United Kingdom
[d] Centre for Chronic Disease Control, New Delhi, India
[e] Madras Diabetes Research Foundation, Chennai, India
[f] Rollins School of Public Health, Emory University, Atlanta, USA
[g] All India Institute of Medical Sciences, New Delhi, India

ARTICLE INFO

Article history:
Received 12 August 2021
Received in revised form
20 October 2021
Accepted 11 November 2021
Available online 16 November 2021

Keywords:
COVID-19
Cessation
Tobacco
Smoking
India

ABSTRACT

Objectives: The Government of India prohibited the sale of tobacco products during the COVID-19 lockdown to prevent the spread of the SARS-CoV-2 virus. This study assessed the tobacco cessation behaviour and its predictors among adult tobacco users during the initial COVID-19 lockdown period in India.

Methods: A cross-sectional study was conducted with 801 adult tobacco users (both smoking and smokeless tobacco) in two urban metropolitan cities of India over a 2-month period (July to August 2020). The study assessed complete tobacco cessation and quit attempts during the lockdown period. Logistic and negative binomial regression models were used to study the correlates of tobacco cessation and quit attempts, respectively.

Results: In total, 90 (11.3%) tobacco users reported that they had quit using tobacco after the COVID-19 lockdown period. Overall, a median of two quit attempts (interquartile range 0–6) was made by tobacco users. Participants with good knowledge on the harmful effects of tobacco use and COVID-19 were significantly more likely to quit tobacco use (odds ratio [OR] 2.2; 95% confidence interval [CI] 1.2–4.0) and reported more quit attempts (incidence risk ratio 5.7; 95% CI 2.8–11.8) compared to those with poor knowledge. Participants who had access to tobacco products were less likely to quit tobacco use compared to those who had no access (OR 0.3; 95% CI 0.2–0.5].

Conclusions: Access restrictions and correct knowledge on the harmful effects of tobacco use and COVID-19 can play an important role in creating a conducive environment for tobacco cessation among users.

Introduction

The COVID-19 pandemic has presented the world with unprecedented challenges for the 21st century, in addition to excess mortalities.[1] Although there remains a lack of evidence to define the risk of COVID-19 infection among tobacco users,[2] these individuals are at an increased risk of adverse outcomes (i.e. death and severity of COVID-19).[3] Recent evidence suggests that smokers have a higher likelihood of COVID-19 complications, including mortality (odds ratio [OR] 1.91; 95% confidence interval [CI] 1.4–2.6).[4] The act of tobacco smoking involves frequent contact between the fingers and mouth and hence can potentially increase the risk of COVID-19 infection.[5] The use of smokeless tobacco (SLT) products, such as gutkha, khaini, zarda, and paan (betel quid with tobacco), induces salivation and hence increased spitting, which may also increase the spread of the SARS-CoV-2 virus.[6]

Stringent tobacco control measures have been enforced by some countries to help prevent the spread of COVID-19. Several countries from the Eastern Mediterranean Region banned the use of waterpipe in indoor and outdoor public places.[7] Bangladesh suspended the production, supply, marketing and sale of all kinds of tobacco

* Corresponding author. HRIDAY, B-5/94, Safdarjung Enclave, First Floor, New Delhi, 110029, India. Tel.: +919810056631.
E-mail address: monika@hriday-shan.org (M. Arora).

https://doi.org/10.1016/j.puhe.2021.11.010

products;[8] Botswana banned the import and sale of cigarettes and other related products;[9] South Africa restricted the sale of cigarettes, snuff, hookah pipes and e-cigarettes to combat the risks posed by the use of tobacco products during the pandemic.[10] The COVID-19 pandemic has provided an opportunity to study the impact of policy environment on tobacco consumption habits of users. A web-based survey in the United States, conducted during the pandemic, showed that 22.9% of respondents attempted to quit smoking cigarettes to reduce their risk of harm from COVID-19 and one-third of respondents reported an increase in motivation to quit tobacco during the pandemic period.[11]

India enforced a nationwide lockdown on 24 March 2020 to prevent the spread of COVID-19.[12] At the beginning of the lockdown, different state governments issued warnings and advisories against tobacco use and about its interwoven relationship with COVID-19. Subsequently, the Indian Council of Medical Research[13] and the Ministry of Health and Family Welfare, Government of India, each issued advisories to prohibit the use and spitting of tobacco.[14,15] Many states in India also announced bans on tobacco use under the troupe of the Indian Penal Code 1890, Cigarettes and Other Tobacco Products Act 2003 and the Epidemic Diseases Act 1897.[16] In addition, the Ministry of Home Affairs, Government of India, prohibited the sale and use of gutkha and other tobacco products in the country,[17] which created a nationwide conducive environment for tobacco control.

Previous evidence suggests that tobacco control policies, such as restricting the availability or access to tobacco products, limit tobacco use.[18,19] A study by Narotam et al.,[20] with 650 participants enrolled in a tobacco cessation counselling programme before the lockdown, reported the positive impact of public health measures on tobacco use behaviour. However, the study only included participants who were already motivated and enrolled in a cessation programme and did not assess the predictors of cessation and quit attempts among other tobacco users. In this study, we aimed to assess tobacco cessation behaviours and identify predictors of tobacco cessation and quit attempts among adult tobacco users during the late COVID-19 lockdown period in India.

Methods

Study design, setting and participants

A cross-sectional study was conducted in two urban metropolitan cities of India (Delhi and Chennai) over a 2-month study period (July–August 2020). Assuming a large (>1 million) target population, a 50% outcome factor in the population, a 5% margin of error and 95% CI, the minimum required sample size was estimated to be 384–400 in the general population. Therefore, a total sample size of 800 participants (approximately 400 from each city) was estimated using open-source epidemiological statistics (OpenEpi).[21] A list of participants from an existing cohort (**C**entre for c**A**rdiometabolic **R**isk **R**eduction in **S**outh Asia),[22] with a history of any form of tobacco use was prepared and individuals were invited to participate in the present study. Participants aged 25 years or more (irrespective of their sex), using any form of tobacco, who spoke English, Hindi or Tamil, and those who provided consent were enrolled in the study. Tobacco users who had recently quit using tobacco (i.e. in the previous 3 months from the date of survey [i.e. after the initial lockdown]) or who had used tobacco in any form during the previous month were also included in the study. Participants who were suffering from any severe illness, institutionalised, unable to respond to the survey or not willing to provide or record verbal consent were excluded from the study.

The objectives of the study were explained to the study participants, and after obtaining informed consent, a telephone questionnaire was administered. Prior ethical approval for research involving human subjects for this study was obtained from the Centre for Chronic Disease Control's Institutional Ethics Committee (Reference #CCDC_IEC_04_2018).

Study tool

The questionnaire was translated, adapted and modified from the STOP survey[23] for the context of smoking and SLT use in India. The survey tool has previously been used in a longitudinal study in Pakistan to capture and compare tobacco use behaviour among users before and during COVID-19.[23] The survey was translated into regional languages (i.e. Hindi [for participants in Delhi] and Tamil [for participants in Chennai]). The survey included questions on sociodemographic variables, knowledge of the adverse effects of tobacco use during COVID-19, intentions to quit tobacco, number of quit attempts and knowledge of tobacco control policies implemented in India during the lockdown period. The questionnaire was piloted on a subgroup of 20 respondents (from each city) and was subsequently adapted before administering it to the study population. A brief description of the study variables is provided in the supplementary file (Table S1).

Data collection and management

Following the rules of social distancing, the questionnaire was administered by telephone, and a standardised protocol was used for data collection. Informed consent was sought from eligible participants. Verbal consent was audio recorded following the recent Indian Council of Medical Research's revised guidelines for obtaining consent for biomedical and health research during the COVID-19 pandemic.[24] The questionnaire was then administered in the language preferred by the participant (i.e. English, Hindi or Tamil).

Data analyses

Descriptive statistics are presented as frequencies and percentages. The primary outcomes of the study were 'cessation' and 'quit attempts'. Participants were asked the question, '*What best describes you?*' and those selecting the option – '*I have stopped using tobacco*' were categorised as 1 for cessation (otherwise, 0). Participants were then asked the question, '*How many attempts to stop tobacco use have you made in the last 6 months?*', and the answers were recorded as an integer. The 'Quit attempts' was treated as discrete (count) data.

Univariate associations were analysed using Fisher's exact/Chi-squared test as appropriate for categorical variables, whereas the count variables were analysed using Mann–Whitney/Wilcoxon test and Kruskal–Wallis test as appropriate. A P value of <0.05 was considered significant. Cross-tabulations between various sociodemographic characteristics (e.g. gender, city, age, education, employment status), knowledge on the harmful effects of tobacco use and COVID-19, knowledge on legislative decisions (taken by government on tobacco sales and consumption during the national lockdown) and access to tobacco products during the lockdown were studied. Responses to all questions assessing the knowledge of participants were aggregated and thereafter scored anonymously. The correct responses were marked as 1 and incorrect as 0. The maximum score for knowledge on the harmful effects and knowledge on legislative decisions was 5, and the minimum score was 0. The aggregate scores were further categorised as poor (mean − 1 standard deviation [SD]), average (mean −1 SD to mean +1 SD) and good (mean +1 SD).[25]

Because of the overdispersion in the number of quit attempts, its associations with various independent variables were studied using the negative binomial regression model,[26] whereas the logistic regression model was used to study the association of independent variables with cessation. The results of the negative binomial regression models and the logistic regression models are given in incidence risk ratio (IRR) and OR with 95% CI, respectively. Variables with a P value <0.15 in the univariate analysis were retained in the multivariate models.[27] Data were analysed using STATA v.13.1 (StataCorp, LP, TX).

Results

Study participant characteristics

A total of 801 tobacco users participated in the survey, including 444 (55.4%) from Delhi and 357 (44.6%) from Chennai (Table 1). As the survey was conducted via telephone, a disposition table[28] is used to explain the response rates. The gross response rate for the study was 48.4%, the basic response rate was 85.3% and the response rate calculated using the CASRO estimator was 60.9%. The detailed disposition table and response rate calculations are provided in supplementary file (Tables S2 and S3). In total, 305 (38.1%) participants were current cigarette smokers, 195 (24.3%) were bidi smokers and 324 (40.4%) were SLT users. There were 90 (11.3%) tobacco users who reported that they had stopped using tobacco at the time of the survey after the lockdown measures were introduced. Overall, a median of two quit attempts (interquartile range [IQR] 0–6) was made by the tobacco users over the past 6 months. The mean scores for knowledge on the harmful effects and knowledge about legislative decisions in the study population were 2.1 (SD 1.9) and 2.7 (SD 2.1), respectively.

Most participants (90.1%) were men. In total, 56.3% of participants were in the 45–64 years age group, followed by 31.6% in the 25–44 years age group and 12.1% in the ≥65 years age group. More than half of the participants were educated either up to high school (39.4%) and intermediary school (31.1%). Most participants were employed (81.1%), whereas the remaining were students (10.9%), housewives (3.6%), retired (2.0%) or unemployed (2.2%; Table 1).

Table 1
Sociodemographic characteristics of the study population (N = 801).

Sociodemographic characteristics	n (%)
City (n = 801)	
Delhi	444 (55.4)
Chennai	357 (44.6)
Sex (n = 801)	
Male	722 (90.1)
Female	79 (9.9)
Age group (n = 801)	
25–44 years	253 (31.6)
45–64 years	451 (56.3)
≥65 years	97 (12.1)
Education (n = 801)	
Illiterate	80 (9.9)
Professional degree/postgraduate	17 (2.1)
Graduate (BA/BSc/BCom/Diploma)	75 (9.4)
Secondary school intermediary	249 (31.1)
High school (class V to IX)	316 (39.4)
Primary school (up to Class IV)	64 (7.9)
Employment status (n = 801)	
Employed	650 (81.1)
Student	88 (10.9)
Housewife	29 (3.6)
Retired	16 (2.0)
Unemployed	18 (2.2)
Current cigarette smokers (n = 801)	305 (38.1)
Current bidi smokers (n = 798)[a]	195 (24.3)
Current SLT users (n = 800)[b]	324 (40.4)

SLT, smokeless tobacco.
[a] Three missing responses for current bidi smokers.
[b] One missing response for current SLT users.

Univariate association of cessation and quit attempts with sociodemographic variables

Cessation and quit attempts were significantly higher in females (cessation 21.5%; number of quit attempts 6.5 [IQR 2–20]) than males (cessation 10.2%; number of quit attempts 2 [IQR 0–5]). The percentage of participants who quit was higher in Chennai (15.4%) than Delhi (7.9%); however, the median number of quit attempts in the past 6 months was higher in Delhi (2 [IQR 0–7]) than in Chennai (1 [IQR 0–4]). Cessation and quit attempts were predominantly higher in housewives (cessation 27.6%; number of quit attempts 12.5 [IQR 7.5–30]) compared with students, employed or retired participants (Table 2).

Univariate association of cessation and quit attempts with knowledge and access

In the univariate analysis, cessation was greater in participants who had no access to tobacco products during the COVID-19 lockdown (19.0%) compared with those who had access (7.8%). Quit attempts were higher in daily bidi smokers (2 [IQR 0–7]) compared with occasional smokers (0 [IQR 0–3]). Whereas in the case of SLT users, quit attempts were higher in occasional SLT users (2 [IQR 0–10]) than in daily users (1 [IQR 0–4]). Quit attempts were predominantly higher in people with good knowledge of the harmful effects of tobacco use during COVID-19 (4 [IQR 0–16]) than participants with average (1 [IQR 0–4]) or poor knowledge (0 [IQR 0–3]). Similarly, quit attempts were also higher in participants with good knowledge on legislative decisions (2 [IQR 0–7]) compared with participants with either average (1 [IQR 0–5]) or poor knowledge (0 [IQR 0–3]; Table 3).

Correlates of cessation and quit attempts

To further determine the correlates that are significantly associated with cessation and quit attempts, logistic regression and negative binomial regression models were used, respectively. Table 4 shows the adjusted OR, IRR and 95% CI for cessation and quit attempts. The final regression models included 797 and 328 participants for cessation and quit attempts, respectively, with complete cases across all variables.

Participants with good knowledge on the harmful effects of tobacco use and COVID-19 were significantly more likely to cease tobacco use than participants with poor knowledge (OR 2.2; 95% CI 1.2–4.0), whereas participants with average knowledge were 50% less likely to cease tobacco use (OR 0.5; 95% CI 0.3–0.9).

Participants with good (OR 0.4; 95% CI 0.2–0.9) and average (OR 0.5; 95% CI 0.3–0.9) knowledge on legislative decisions were 60% and 50%, respectively, more likely to cease tobacco use than those with poor knowledge on legislative decisions.

Participants who had access to tobacco products were 70% less likely to cease tobacco use (OR 0.3; 95% CI 0.2–0.5) compared with those who had access to tobacco.

Quit attempts were significantly more likely to occur in participants with average (IRR 1.9; 95% CI 1.0–3.4) and good (IRR 5.7; 95% CI 2.8–11.8) knowledge on the harmful effects of tobacco use and COVID-19 compared with participants with poor knowledge. However, no significant associations for quit attempts were observed among participants with average or good knowledge on legislative decisions.

Table 2
Univariate association of cessation and quit attempts with sociodemographic characteristics.

Variables	Cessation[a]		Quit attempts[b]	
	n (%)	*P* value	Median (IQR)	*P* value
Overall	90 (11.3)		2 (0–6)	
Gender		**0.002***		**0.001$**
Male ($n = 718$)	73 (10.2)		2 (0–5)	
Female ($n = 79$)	17 (21.5)		6.5 (2–20)	
City		**0.001***		**0.024$**
Chennai ($n = 357$)	55 (15.4)		1 (0–4)	
Delhi ($n = 440$)	35 (7.9)		2 (0–7)	
Age group		0.818		0.107
25–44 years ($n = 253$)	26 (10.3)		2 (0–7)	
45–64 years ($n = 447$)	53 (11.9)		1 (0–5)	
≥65 years ($n = 97$)	11 (11.3)		5 (0–4)	
Education		0.162		0.531
Illiterate ($n = 80$)	7 (8.7)		1 (0–10)	
Professional degree/postgraduate ($n = 16$)	5 (31.2)		4 (2–10)	
Graduate ($n = 74$)	7 (9.5)		2 (0–7)	
Secondary/intermediary schools ($n = 248$)	26 (10.5)			
High school ($n = 315$)	39 (12.4)		2 (0–5)	
Primary schools ($n = 64$)	6 (9.4)		0 (0–2)	
Employment status		0.102		**0.020&**
Employed ($n = 648$)	68 (10.5)		2 (0–5.5)	
Student ($n = 86$)	10 (11.6)		2 (0–6.5)	
Housewife ($n = 29$)	8 (27.6)		12.5 (7.5–30)	
Retired ($n = 16$)	2 (12.5)		1 (0–2)	
Unemployed ($n = 18$)	2 (11.3)		1 (0–3)	

Bold *P*-values indicate significant association.
IQR, interquartile range.
*P value <0.05 using Chi-squared test.
$$P$ value <0.05 using Mann–Whitney Wilcoxon test.
&P value <0.05 using Kruskal–Wallis H test.
[a] 4 missing responses for cessation.
[b] Total 329 responses for quit attempts.

Discussion

The COVID-19 pandemic provided a unique opportunity for the promotion of tobacco control strategies, nationally as well as globally.[29,30] The tobacco control policies implemented to address the spread of COVID-19, including restricting access to tobacco products, led to favourable circumstances for tobacco cessation among users.[31] In the present study, 11.3% of tobacco users stopped using tobacco during the lockdown. On average, two quit attempts were made by tobacco users during the past six months. The percentage of people who ceased tobacco use was much lower than reported in a previous study (51%) by Gupte et al.[20] However, this may be attributed to the fact that the population in the study by Gupte et al. comprised of individuals who were already enrolled in a tobacco cessation programme, thus were already motivated to quit tobacco use. On the contrary, some studies from other countries have reported an increase in smoking during the pandemic because of high levels of stress and boredom.[32,33]

Existing evidence suggests that there are low levels of knowledge on the harmful effects of tobacco use and COVID-19.[34,35] The results of the present study also show low levels of knowledge on the harmful effects of tobacco use and COVID-19 among study participants during the lockdown period. Despite this, the results suggest that participants with good knowledge on the harmful effects of tobacco use and COVID-19 were more likely to cease tobacco use and make attempts to quit compared with those with poor knowledge. These findings are consistent with those of a previous study conducted in India[20] and indicate that good knowledge on the harmful effects of tobacco use and COVID-19 could discourage tobacco use among existing users. Moreover, good knowledge on the legislative decisions also seemed to motivate tobacco cessation among users.

Technology has played a vital role in enabling routine and professional activities to continue during the pandemic.[36] Hence, cessation efforts (e.g. creating awareness of tobacco use during COVID-19 and cessation services) via digital media (e.g. television, internet and social media) can be useful.[37,38] Information Communication and Technology can help in propelling and strengthening tobacco control policies.[39] Informative advertising (e.g. harmful effects of tobacco use during COVID-19, knowledge about the National Quitline and m-cessation services) in local languages and dialects can further motivate the cessation of tobacco use.[29] These advertisements should be comprehensively integrated with other commonly used digital applications or social media websites to create awareness among tobacco users.[29]

In the present study, cessation was more prevalent in tobacco users who had no access to tobacco products (19.0%). In fact, cessation was 70% less likely among participants reporting access to tobacco products. The national lockdown during the early months of the COVID-19 pandemic curbed access to tobacco products and may have encouraged abstinence from tobacco among existing users. However, quitting tobacco is often associated with a high relapse rate,[40,41] and there is a high likelihood that users who reported having ceased tobacco use during this period might subsequently relapse after the end of lockdown restrictions. Implementing non-price-based tobacco control policies (e.g. tobacco use restrictions in working places, restriction on access to tobacco products) is considered to be a highly cost-effective measure.[42] Therefore, the ban on the sale of tobacco products and spitting in public places, in addition to designating these acts as an

Table 3
Univariate association of cessation and quit attempts with knowledge and accessibility of tobacco products.

Variables	Cessation[a]		Quit attempts[b]	
	n (%)	*P* value	Median (IQR)	*P* value
Overall	90 (11.3)		2 (0–6)	
Cigarette smoking				
Current daily cigarette smokers ($n = 220$)	NA		1 (0–5)	0.107
Current occasional smokers ($n = 85$)	NA		1 (0–5)	
Non-smokers ($n = 496$)	NA		2 (0–7)	
Bidi smokers				**0.002****
Current daily bidi smokers ($n = 167$)	NA		2 (0–7)	
Current occasional bidi smokers ($n = 28$)	NA		0 (0–3)	
Non-smokers ($n = 603$)			0 (0–10)	
SLT users				**0.000****
Current daily SLT users ($n = 266$)	NA		1 (0–4)	
Current occasional SLT users ($n = 58$)	NA		2 (0–10)	
Non-users ($n = 476$)			5 (1–60)	
Knowledge on the harmful effects				
Poor ($n = 267$)	41 (15.4)	**0.000***	0 (0–3)	**0.000****
Average ($n = 298$)	16 (5.4)		1 (0–4)	
Good ($n = 232$)	33 (14.2)		4 (0–16)	
Knowledge on legislative decisions				
Poor ($n = 231$)	45 (19.5)	**0.000***	0 (0–3)	**0.018****
Average ($n = 280$)	25 (8.9)		1 (0–5)	
Good ($n = 286$)	20 (6.9)		2 (0–7)	
Accessibility of tobacco products				
Yes ($n = 550$)	43 (7.8)	**0.000***	2 (0–6)	0.649
No ($n = 47$)	47 (19.0)		1 (0–4)	

Bold *P*-values indicate significant association.
IQR, interquartile range; SLT, smokeless tobacco.
*P value <0.05 using Fisher's chi exact test.
**P value <0.05 using Kruskal–Wallis H test.
[a] Four missing responses for cessation.
[b] Total 329 responses for quit attempts.

offence with huge penalties for violations, should be considered a public health strategy, both to overcome the COVID-19 pandemic and for tobacco control in India.[43,44] Continuation of the tobacco ban can be justified, as the COVID-19 pandemic is far from over; however, the restrictions laid down by the government (i.e. limiting access to tobacco products) should be monitored closely. Restricting access to tobacco products requires multisectoral regulatory policies and a whole-society approach so that users can be supported to quit and initial uptake can be prevented.

The COVID-19 pandemic provides a conducive policy environment to implement tobacco control strategies to reduce the production as well as consumption of tobacco products. Implementing demand-reduction strategies,[45] such as the ban on tobacco use and spitting in public places, and raising awareness of the harmful effects of tobacco use during COVID-19 can further strengthen tobacco control policies. Similarly, curtailing tobacco supply,[45] by limiting the access to products, can further help address both the COVID-19 pandemic and tobacco epidemic. The results from this study can be used to align population- and individual-level interventions, including drawing on national-level change to encourage greater participation in tobacco cessation programmes. Sustained efforts may help substantially reduce tobacco use, with the possibility of eliminating tobacco use in the future.

Strengths and limitations

The present study enrolled participants from an existing cohort.[22] Respondents were followed up during the COVID-19 pandemic to assess the impact of COVID-19 restrictions on tobacco use cessation and quit attempts. A previous study investigated tobacco cessation behaviour during COVID-19 lockdown in participants enrolled in a tobacco-cessation programme; thus, this study population was already motivated to cease tobacco use.[20] Participants in the present study were not premotivated to cease tobacco use, and hence, their behaviour can be attributed to the pandemic alone.

This study attempted to investigate cessation and quit attempts among tobacco users during the COVID-19 crisis, but there are limitations to this study. Cessation is generally defined as the abstinence from tobacco use for a minimum period of 6–12 months. However, as this was a rapid study conducted over a 2-month period (during the COVID lockdown period in India), patients who reported that they had stopped using tobacco completely since the start of the lockdown were considered to have ceased tobacco use. The number of quit attempts was reported over 6 months; the survey was conducted in the months of July and August, but the number of quit attempts could also include attempts made before the study period. This study presents estimates based on a single study conducted in two large Indian cities (Delhi and Chennai). Furthermore, the cohort was limited to urban areas of the country and does not include tobacco users aged <25 years. Therefore, the findings of the study cannot be generalised to all tobacco users in India. Thus, we recommend that large population-based interstate studies are used to further evaluate the effects of restrictions in access to tobacco products on tobacco use cessation.

Conclusion

Measures enforced by the Government of India to reduce access to tobacco products during the nationwide COVID-19 lockdown led to a favourable environment for existing tobacco users to quit. This highlights an opportunity to align communicable and non-communicable disease responses during a public health crisis and could provide lessons for future tobacco control efforts. The m-

Table 4
Correlates of cessation and quit attempts.

Variables	Cessation [OR (95% CI)][a] $n = 797$	Quit attempts [IRR (95% CI)][b] $n = 328$
Gender		
Male	Ref	Ref
Female	1.3 (0.5–3.1)	1.9 (0.5–6.7)
City		
Chennai	Ref	Ref
Delhi	0.6 (0.3–1.1)	0.6 (0.3–1.1)
Age group		
25–44 years	NA	Ref
45–64 years	NA	0.9 (0.6–1.5)
≥65 years	NA	0.7 (0.2–2.1)
Employment status		
Employed	Ref	Ref
Student	1.0 (0.4–2.4)	1.1 (0.3–3.5)
Housewife	1.9 (0.6–6.1)	0.9 (0.2–5.4)
Retired	1.5 (0.3–7.0)	0.2 (0.0–1.0)
Unemployed	0.8 (0.1–3.9)	0.3 (0.1–2.0)
Cigarette smokers		
Non-users	NA	Ref
Current daily cigarette smokers	NA	0.7 (0.3–1.8)
Current occasional smokers	NA	0.7 (0.3–1.6)
Bidi smokers		
Non-users	NA	Ref
Current daily bidi smokers	NA	**0.2 (0.1–0.6)**
Current occasional bidi smokers	NA	0.7 (0.2–2.2)
SLT users		
Non-users	NA	Ref
Current daily SLT users	NA	0.7 (0.2–1.7)
Current occasional SLT users	NA	1.4 (0.4–4.5)
Knowledge on the harmful effects of tobacco use and COVID-19		
Poor	Ref	Ref
Average	**0.5 (0.3–1.0)**	**1.9 (1.0–3.4)**
Good	**2.2 (1.2–4.0)**	**5.7 (2.8–11.8)**
Knowledge on legislative decisions		
Poor	Ref	Ref
Average	**0.5 (0.3–0.9)**	1.5 (0.8–2.9)
Good	**0.4 (0.2–0.9)**	1.6 (0.7–3.6)
Overall access		
No	Ref	NA
Yes	**0.3 (0.2–0.5)**	NA

CI, confidence interval; IRR, incidence risk ratio; OR, odds ratio; Ref, reference; SLT, smokeless tobacco.
Bold values indicate significant association.
[a] Estimates derived using logistic regression. Variables with P values <0.15 in univariate analysis were included in the regression models.
[b] Estimates derived using negative binomial regression model. Variables with P values <0.15 in univariate analysis were included in the regression models.

cessation, Quitline and in-person cessation services should be provided proactively during this opportune time to encompass more tobacco users and help encourage cessation and quit attempts.

Author statements

Acknowledgements

The authors would like to acknowledge the data collection teams from HRIDAY, New Delhi and Chennai for participating in the study. The authors would also like to acknowledge the CARRS team for providing the list of participants and relevant information from the last follow-up.

Ethical approval

Ethical approval for this study was obtained from the Centre for Chronic Disease Control's Institutional Ethics Committee (Reference #CCDC_IEC_04_2018).

Funding

The activities and results presented in this article were supported by the following grants: the COVID-19 and Tobacco Project, funded from the University of Edinburgh's Scottish Funding Council Global Challenges Research Fund (GCRF) to the Tobacco Control Capacity Programme (MR/P027946/2), supported by UK Research and Innovation (UKRI) with funding from the Global Challenges Research Fund.

Competing interests

None declared.

Authors' contributions

M.A. and G.P.N. conceptualised the study. L.B. and M.A. secured funding for the study. M.A., G.P.N. and N.J. adapted the study tool. G.P.N., N.S. and N.J. facilitated the data collection and implementation of the study. N.S. and G.P.N. analysed the results, and all the authors contributed to the interpretation of the findings. M.A., G.P.N., N.S. and N.J. drafted the article. S.M., D.P., D.M., L.B., K.S.R., M.K.A., V.M., N.T., K.M.V.N. and F.D. revised the article. All authors approved the final article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2021.11.010.

References

1. World Health Organization Coronavirus Disease (COVID-19) Dashboard. Accessed November 27, 2020. https://covid19.who.int/.
2. Smoking and COVID-19. Accessed November 27, 2020. https://www.who.int/news-room/commentaries/detail/smoking-and-covid-19.
3. World Health Organization Statement: Tobacco use and COVID-19. Accessed November 27, 2020. https://www.who.int/news/item/11-05-2020-who-statement-tobacco-use-and-covid-19.
4. Patanavanich R, Glantz SA. *Smoking is associated with COVID-19 progression: a meta-analysis nicotine & tobacco research*. Published online. 2020.
5. El-Awa F, Fraser CP, Adib K, et al. The necessity of continuing to ban tobacco use in public places post-covid-19. *East Mediterr Health J* 2020;**26**(6):630–2. https://doi.org/10.26719/emhj.20.071.
6. World Health Organization. Coronavirus disease (COVID-19): tobacco. Accessed November 27, 2020. https://www.who.int/news-room/q-a-detail/coronavirus-disease-covid-19-tobacco.
7. World Health Organization. *Advancing in tobacco control amidst the COVID-19 pandemic*. Published online. 2020. http://www.who.int/fctc/mediacentre/news/2020/advancing-in-tobacco-control-amidst-the-covid-19-pandemic/en/. [Accessed 27 November 2020].
8. Bangladesh orders to stop production, marketing, sale of tobacco products - national - observerbd.com. Accessed November 27, 2020. https://www.observerbd.com/news.php?id=257281.
9. COVID-19: Botswana bans imports, sales of cigareottes, related products - Apanews.net. Accessed November 27, 2020. http://apanews.net/en/news/covid-19-botswana-bans-imports-sales-of-cigarettes-related-products/.
10. Egbe CO, Ngobese SP. COVID-19 lockdown and the tobacco product ban in South Africa. *Tob Induc Dis* 2020;**18**(May). https://doi.org/10.18332/TID/120938.
11. Klemperer EM, West JC, Peasley-Miklus C, Villanti AC. Change in tobacco and electronic cigarette use and motivation to quit in response to COVID-19. *Nicotine Tob Res* 2020;**22**(9):1662–3.
12. Coronavirus India lockdown day 68 updates. May 31, 2020 - The Hindu. Accessed November 27, 2020. https://www.thehindu.com/news/national/india-coronavirus-lockdown-may-31-2020-live-updates/article31714194.ece.
13. ICMR appeals to people to not chew, spit tobacco products in public. Headlines. Accessed November 27, 2020. https://www.devdiscourse.com/article/headlines/993152-icmr-appeals-to-people-to-not-chew-spit-tobacco-products-in-public.
14. 28 states, UTs ban smokeless tobacco products, spitting due to coronavirus - The Economic Times. Accessed November 27, 2020. https://economictimes.indiatimes.com/news/politics-and-nation/28-states-uts-ban-smokeless-tobacco-products-spitting-due-to-coronavirus/articleshow/75657382.cms?from=mdr.
15. COVID-19 pandemic and tobacco use in India. doi:10.4103/0019-509X.147477.
16. Lal P, Mishra D, Singh R. A "Bottom-up approach" to introduce ban on tobacco products to prevent spitting during COVID-19: an early review of progress made and challenges. *Int J Noncommunicable Dis* 2020;**5**(3). 138-138, https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=24688827&v=2.1&it=r&id=GALE%7CA637698957&sid=googleScholar&linkaccess=fulltext. [Accessed 27 November 2020].
17. Consolidated revised guidelines on the measures to be taken by Ministries, Departments of Government of India, state, UT/governments and state/UT authorities for containment of COVID-19 in country. Accessed November 27, 2020. https://prsindia.org/files/covid19/notifications/2732.IND_Revised_Consolidated_Guidelines_April_15.pdf.
18. Méndez D, Alshanqeety O, Warner KE. The potential impact of smoking control policies on future global smoking trends. *Tob Control* 2013;**22**(1):46–51. https://doi.org/10.1136/tobaccocontrol-2011-050147.
19. Levy DT, Tam J, Kuo C, Fong GT, Chaloupka F. The impact of implementing tobacco control policies: the 2017 tobacco control policy scorecard. *J Public Health Manag Pract* 2018;**24**(5):448–57. https://doi.org/10.1097/PHH.0000000000000780.
20. Gupte H, Mandal G, Jagiasi D. How has the COVID-19 pandemic affected tobacco users in India: lessons from an ongoing tobacco cessation program. *Tob Prev Cessat* 2020;**6**(September):1–5. https://doi.org/10.18332/tpc/127122.
21. OpenEpi Menu. Accessed November 27, 2020. http://www.openepi.com/Menu/OE_Menu.htm.
22. Nair M, Ali MK, Ajay VS, et al. CARRS Surveillance study: design and methods to assess burdens from multiple perspectives. *BMC Public Health* 2012;**12**(1):701. https://doi.org/10.1186/1471-2458-12-701.
23. STOP survey (Studying Tobacco users of Pakistan (STOP): repeated cross-sectional and longitudinal survey) – The Initiative. Accessed November 27, 2020. http://theinitiative.org.pk/stop-survey-studying-tobacco-users-of-pakistan-stop-repeated-cross-sectional-and-longitudinal-survey/.
24. National Guidelines for Ethics Committees Reviewing Biomedical & Health Research during COVID-19 pandemic. Accessed November 27, 2020. https://www.icmr.gov.in/pdf/covid/techdoc/EC_Guidance_COVID19_06052020.pdf.
25. Ali M, Uddin Z, Banik PC, et al. *Knowledge, attitude, practice and fear of COVID-19: a cross-cultural study*. Published online. 2020. p. 1–27. https://doi.org/10.1101/2020.05.26.20113233.
26. Byers AL, Allore H, Gill TM, Peduzzi PN. Application of negative binomial modeling for discrete outcomes: a case study in aging research. *J Clin Epidemiol* 2003;**56**(6):559–64. https://doi.org/10.1016/S0895-4356(03)00028-3.
27. Bursac Z, Gauss CH, Williams DK, Hosmer DW. Purposeful selection of variables in logistic regression. *Source Code Biol Med* 2008;**3**(1):17. https://doi.org/10.1186/1751-0473-3-17.
28. Gripp SI, Luloff AE, Yonkers RD. Reporting response rates for telephone surveys used in agricultural economics research. *Agric Resour Econ Rev* 1994;**23**:200–6 (1203-2016-95023).
29. Grover S, Mohanty V, Jain S, Anand T, Aghi MB. "YES it's the perfect time to quit": fueling tobacco cessation in India during COVID-19 pandemic. *Tob Use Insights* 2020;**13**. https://doi.org/10.1177/1179173x20960447. 1179173X2096044.
30. Kayhan Tetik B, Gedik Tekinemre I, Taş S. The effect of the COVID-19 pandemic on smoking cessation success. *J Community Health* 2020:1. https://doi.org/10.1007/s10900-020-00880-2. Published online.
31. Ahluwalia IB, Myers M, Cohen JE. COVID-19 pandemic: an opportunity for tobacco use cessation. *Lancet Public Health* 2020;**5**(11):e577. https://doi.org/10.1016/S2468-2667(20)30236-X.
32. Bommele J, Hopman P, Walters BH, et al. The double-edged relationship between COVID-19 stress and smoking: implications for smoking cessation. *Tob Induc Dis* 2020:18.
33. Sidor A, Rzymski P. Dietary choices and habits during COVID-19 lockdown: experience from Poland. *Nutrients* 2020;**12**(6):1657.
34. Yach D. Tobacco use patterns in five countries during the COVID-19 lockdown. *Nicotine Tob Res* 2020;**22**(9):1671–2. https://doi.org/10.1093/ntr/ntaa097.
35. Kalan ME, Ghobadi H, Taleb Z Ben, et al. COVID-19 and beliefs about tobacco use: an online cross-sectional study in Iran. *Environ Sci Pollut Res* 2020:1–9. https://doi.org/10.1007/s11356-020-11038-x. Published online October 7.
36. Whitelaw S, Mamas MA, Topol E, Van Spall HGC. Applications of digital technology in COVID-19 pandemic planning and response. *Lancet Digit Health* 2020. Published online.
37. Morriscey C, Shephard A, van Houdt A, Kerr D, Barrett SP. Using 'Smart' Technology to aid in cigarette smoking cessation: examining an innovative way to monitor and improve quit attempt outcomes. *J Smok Cessat* 2019;**14**(3):149–54.
38. World Health Organization. The missing O: scaling up tobacco cessation in India using mobile technology. Accessed March 18, 2021. https://www.who.int/tobacco/quitting/scaling-up-cessation-india-using-mobile-technology/en/.
39. Mahoney MC, Erwin DO, Twarozek AM, et al. Leveraging technology to promote smoking cessation in urban and rural primary care medical offices. *Prev Med (Baltim)* 2018;**114**:102–6.
40. García-Rodríguez O, Secades-Villa R, Flórez-Salamanca L, Okuda M, Liu S-M, Blanco C. Probability and predictors of relapse to smoking: results of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). *Drug Alcohol Depend* 2013;**132**(3):479–85.
41. Edwards KC, Kasza KA, Tang Z, et al. Correlates of tobacco product reuptake and relapse among youth and adults in the USA: findings from the PATH Study Waves 1–3 (2013–2016). *Tobac Control* 2020;**29**(Suppl 3):s216–26.
42. Ekpu VU, Brown AK. The economic impact of smoking and of reducing smoking prevalence: review of evidence. *Tob Use Insights* 2015;**8**. TUI-S15628.
43. Policy responses to smokeless tobacco (ST) in India during the COVID19 pandemic | ORF. Accessed March 18, 2021. https://www.orfonline.org/expert-speak/policy-responses-to-smokeless-tobacco-st-in-india-during-the-covid19-pandemic/.
44. India's regulatory struggle to rein in smokeless tobacco use: is Covid-19 a game changer? | Blog - Tobacco Control. Accessed March 18, 2021. https://blogs.bmj.com/tc/2020/06/28/indias-regulatory-struggle-to-rein-in-smokeless-tobacco-use-is-covid-19-a-game-changer/.
45. Warner KE. *An endgame for tobacco?*. 2013. Published online.

Public Health 202 (2022) 80–83

Contents lists available at ScienceDirect

Public Health

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

Short Communication

COVID-19 incidence in border regions: spatiotemporal patterns and border control measures

Tobias Chilla [a], Tim Große [a], Stefan Hippe [a], Blake Byron Walker [*, a]

Friedrich-Alexander University, Erlangen, Germany

ARTICLE INFO

Article history:
Received 8 September 2021
Received in revised form
22 October 2021
Accepted 4 November 2021
Available online 15 November 2021

Keywords:
Public health
COVID-19
Borders
Epidemiology
Geography

ABSTRACT

Objectives: Among the few studies examining patterns of COVID-19 spread in border regions, findings are highly varied and partially contradictory. This study presents empirical results on the spatial and temporal dynamics of incidence in 10 European border regions. We identify geographical differences in incidence between border regions and inland regions, and we provide a heuristic to characterise spillover effects.
Study design: Observational spatiotemporal analysis.
Methods: Using 14-day incidence rates (04/2020 to 25/2021) for border regions around Germany, we delineate three pandemic 'waves' by the dates with the lowest recorded rates between peak incidence. We mapped COVID-19 incidence data at the finest spatial scale available and compared border regions' incidence rates and trends to their nationwide values. The observed spatial and temporal patterns are then compared to the time and duration of border controls in the study area.
Results: We observed both symmetry and asymmetry of incidence rates within border pairs, varying by country. Several asymmetrical border pairs feature temporal convergence, which is a plausible indicator for spillover dynamics. We thus derived a *border incidence typology* to characterise (1) symmetric border pairs, (2) asymmetric border pairs without spillover effects, and (3) asymmetric with spillover effects. In all groups, border control measures were enacted but appear to have been effective only in certain cases.
Conclusions: The heuristic of border pairs provides a useful typology for highlighting combinations of spillover effects and border controls. We conclude that border control measures may only be effective if the timing and the combination with other non-pharmaceutical measures is appropriate.

Introduction

Efforts to stem the interregional spread of SARS-CoV-2 rely largely on carefully timed policies to reduce vector mobility.[1] Border management has therefore been a highly visible and often utilised form of non-pharmaceutical pandemic control following the SARS-CoV-2 outbreak, despite the rather controversial and legally complex nature of border policy.[2–4]

Very few studies have examined patterns of COVID-19 spread and containment in and across border regions, with several notable exceptions. Alizon et al.[5] and Klatt[6] conducted regional analyses of COVID-19, providing some initial data, while Duvernet[7] found that border control measures do not show a clear impact in the German context. Conversely, Hossain et al.[8] postulate a positive effect in mainland China. Scarpone et al.[9] identified clusters of high COVID-19 incidence rates in several of Germany's border regions and underscored the need for further research to directly examine the role of borders and border controls in SARS-CoV-2 containment.

This report responds directly to this need for closer investigation by providing initial results from our examination of border regions' roles in the spatial and temporal dynamics of COVID-19, focussing on Germany and its nine bordering countries. The objectives of this study were to (i) identify geographical differences in incidence rates between border regions and inland regions in the study area, (ii) identify possible spill-over effects in border regions, and (iii) conduct a preliminary examination of the spatiotemporal association between border controls and COVID-19 incidence.

Methods

This analysis uses 14-day incidence rates (reported confirmed cases per 100,000 inhabitants) from week 04/2020 to week 25/2021 for border regions in and around Germany. Data were acquired from open data sources at the European Centre for Disease Prevention and Control and the Swiss Federal Office of Public Health. For several

* Corresponding author. Wetterkreuz 15, 91058 Erlangen, Germany. Tel.: +49 157 3419 1735.
E-mail address: blake.walker@fau.de (B.B. Walker).
[a] These authors contributed equally to this article and share the first authorship.

https://doi.org/10.1016/j.puhe.2021.11.006

regions, data were only available beginning in week 18/2020. We delineated three pandemic 'waves', differentiated by the dates with lowest recorded rates between peak incidence in Germany and the results were mapped and analysed. As early-pandemic data were only available for Belgium, Germany, and France, we focus particularly on waves 2 and 3, which feature complete data for all ten countries.

Our analysis uses COVID-19 incidence data at the finest spatial scale for which data were available (NUTS3 in DE, CZ; NUTS2 in PL, NL, CH, FR, DK, AT; NUTS1 in BE), and we compare their border regions' incidence rates to their nationwide values (NUTS0). Administrative areas adjacent to a national border (including NUTS3 regions with min. 25% of their area within a 25 km buffer along the border, see Supplement S1) were selected and assigned a categorical variable for pandemic wave and country. The border regions' incidence rates presented herein result from the non-weighted averages of their subunits.

(Land) Border control measures were defined as official suspensions of the Schengen agreement.[10] The incidence rates and border control measures were analysed on the level of border pairs. The German-Czech border region was divided into two border pairs due to large differences in the spatiotemporal patterns between the CZ-Bavaria (BY) and CZ-Saxony (SN) regions. Bivariate linear models were run to examine the strength of correlation for incidence rates between border pairs.

Results

Many border pairs exhibit differences in their spatial and temporal development (see Supplement S1). These results exhibit both symmetry and asymmetry of incidence across border pairs. Several asymmetrical border pairs feature temporal convergence, which is a plausible indicator for spillover dynamics.

Based on the observed spatiotemporal patterns, we derived the following *border incidence typology,* summarising both a/symmetry and convergence/spill over. Four representative border pairs are shown in Fig. 1, and all border pairs can be found in the appendix (Supplement S2–7).

(1) **Symmetric border pairs** feature similar incidence rates and/or congruent trends on both sides of the border. In these instances, the border does not appear to inhibit infection dynamics. The border pair DK-DE in the first wave is a typical example featuring parallel temporal development (Fig. 1).
(2) **Asymmetric border pairs without spillover effects** exhibit clear differences in their incidence rates and/or temporal trends. Over time, the values do not converge, i.e., spillover effects across the border appear unlikely. The BE-DE pair during the second wave shows clearly contrasting trends (a ten-fold difference, see Fig. 1). During this time, German border regions feature similar patterns as the German inland rates. As such, this asymmetry may indicate effective containment inhibiting any spillover effects.
(3) Several border pairs are **asymmetric with spillover effects,** showing clear differences in the level of incidence while featuring temporal convergence. In these cases, a time lag between the peak incidence rates and an incomplete convergence are visible. The border pair CZ-BY during the second wave illustrates this pattern, with the incidence in Bavaria rising 8–9 weeks after the initial spike in the Czech Republic. During this time, the incidence rates in the German border region are notably higher than the national average, suggesting a spillover effect in this region.

While the CZ-BY case displays moderate evidence of spill over, the CZ-SN border pair during the second wave case features a much stronger trend. The time lag is also approximately 8–9 weeks, but the convergence effect is much stronger (Fig. 1). The temporal dynamics in this region may provide an indication of a recurrent 'yo-yo' effect throughout the second and third waves. The incidence rates in Saxony appear to be decoupled from the German average values, which may provide further evidence of a spillover effect (in contrast to Duvernet,[7] but similar to Hossain et al.[8]).

For the asymmetric border pairs with spillover effects, bivariate linear regressions indicated a high degree of correlation between incidence rates on either side of the border. For example, in the second wave, the CZ-SN pair exhibits a high degree of similarity with an R-square value of 0.82 ($P < 0.001$), while CZ-BY has an R-square of 0.666 ($P < 0.001$). Other border pairs featured a range between 0.451 and 0.262 ($P < 0.073$). These preliminary results provide evidence of a statistical dependence in all asymmetric border pairs with spillover effects. Asymmetrical border pairs without spillover effects do not feature correlation of rates.

All ten border pairs are therefore categorised by their a/symmetry and duration of border control measures. Among those with no border controls (or brief controls lasting less than 14 days), PL-DE, CZ-BY, and CZ-SN all exhibited asymmetry with spillover effects in the second wave. For the CZ-SN border pair, the difference between the second and third waves is strongly pronounced. In contrast, the border controls throughout the third wave are coincident with a weaker degree of symmetry. The remaining border pairs during waves 1 and 3 did not exhibit observable cross-border spillover effects.

Among borders that experienced border-crossing restrictions of 14 days or longer, symmetrical spatiotemporal development was observed in the first wave for DK-DE, PL-DE, and CZ-SN, and recurrent symmetry for DK-DE in the third wave. Also, during the third wave, CZ-SN and CZ-BY experienced border controls but still exhibited clear spillover effects. Border controls in other regions coincide with a lack of observable spill over, suggesting that the border-crossing restrictions may have been effective.

A notable outlier was observed for the Belgian-German border. During the peak incidence period of the second wave, no border controls were enacted. Nevertheless, a high degree of asymmetry suggests minimal or no spillover effects. In contrast, the third wave was accompanied by border controls, corresponding with increasingly asymmetrical development.

Discussion

Our preliminary results indicate that national borders may play a role in explaining the observed patterns of COVID-19 incidence, despite being within the Schengen Area. The outliers noted in our observations suggest that border controls were not universally effective for preventing spillover effects and that their effectiveness may be more closely related to the specific restrictions and means of enforcement, underscoring the need for detailed case studies to ascertain specifically the categories of measures that may be effective.

While some borders exhibited strong cross-border spillover effects, others appear to have been effective at controlling the spread of SARS-CoV-2. However, the ability to infer the effect of border controls may be inhibited by missing incidence data for certain regions in the early phases of the pandemic, as well as different spatial resolutions, i.e., that some regions report incidence by NUTS2 whil others report at the NUTS3 scale. Border control measures are one of many non-pharmaceutical measures, such as lock-ins, compulsory masks, and social distancing; the observed cross-border differences may reflect policy differences in this regard. However, the complexity of the observed patterns underscores the need for detailed examination of specific border control measures and consideration of how multiple non-pharmaceutical measures can be utilised.

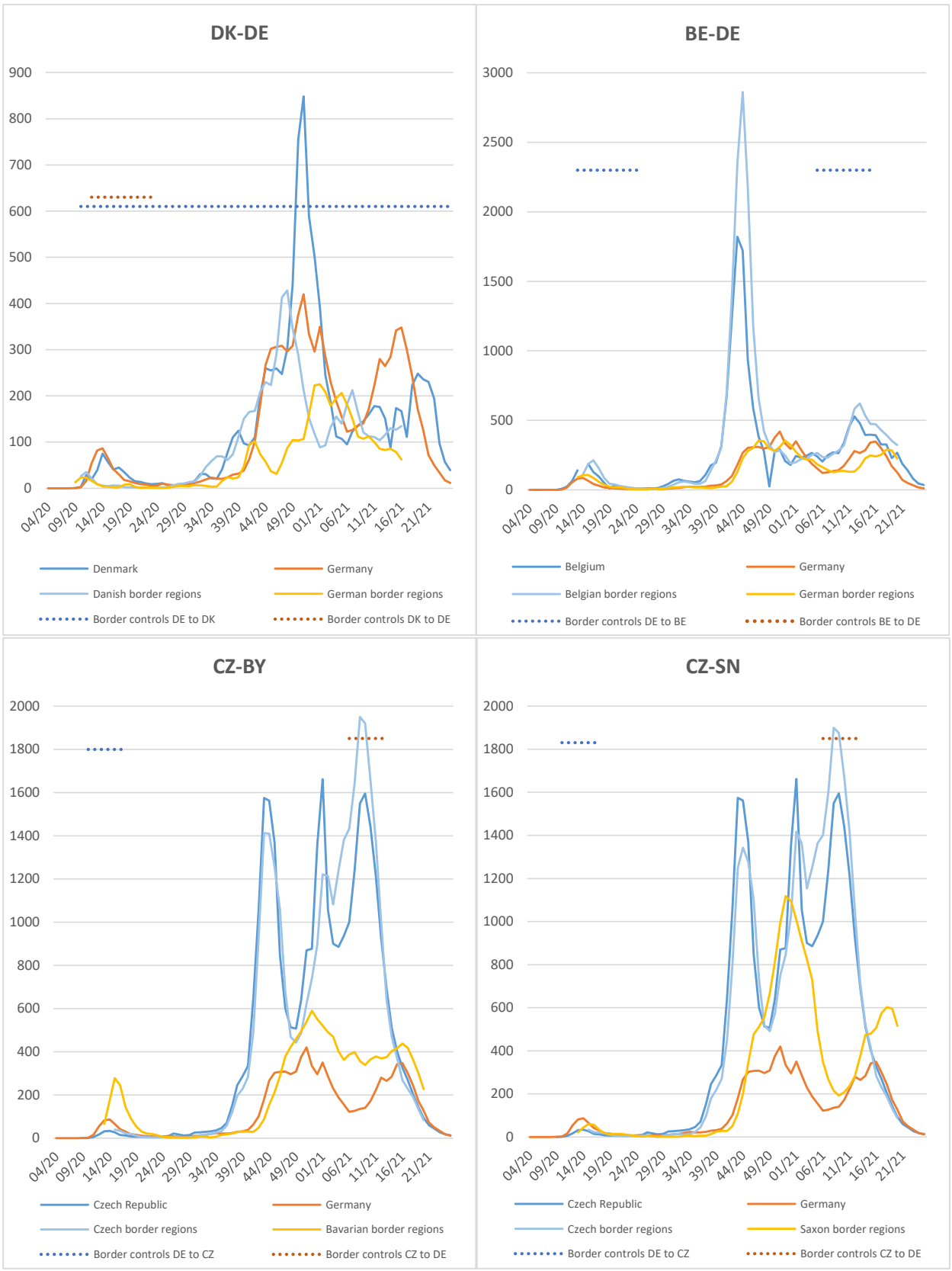


Fig. 1. Timeline of incidence rates and border controls for four selected border pairs. Axes labeles: Y = Confirmed cases per 100,000 inhabitants, X = Week.

Author statements

Acknowledgements

The authors wish to extend their thanks to all data providers listed herein.

Ethical approval

This analysis only uses anonymised, open data at the aggregate level and is therefore low-risk and not subject to ethics approvals at the host institution.

Funding

This study did not receive any external funding. T.G. was supported by supplemental research funding provided by Friedrich-Alexander-Universität Erlangen-Nürnberg. B.B.W. is supported by the German Federal Ministry for Education and Research (BMBF).

Competing interests

The authors declare that they have no conflict of interest in the submission and publication of this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2021.11.006.

References

1. Flaxmann S, Mishra S, Gandy A, Unwin JT, Mellan TA, Coupland H, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* 2020;**584**:257–61. https://doi.org/10.1038/s41586-020-2405-7.
2. Opiłowska E. The Covid-19 crisis: the end of a borderless Europe? *Eur Soc* 2021;**23**:589–600. https://doi.org/10.1080/14616696.2020.1833065.
3. Weber F, Wille C. Grenzgeographien der COVID-19-Pandemie. In: Weber F, Wille C, Caesar B, Hollstegge J, editors. *Geographien der Grenzen. Räume – Grenzen – Hybriditäten*. Wiesbaden: Springer VS; 2020. p. 191–223. https://doi.org/10.1007/978-3-658-30950-3_9.
4. Zhu Z, Weber E, Strohsal T, Serhan D. Sustainable border control policy in the COVID-19 pandemic: a math modeling study. *Trav Med Infect Dis* 2020;**41**:102044. https://doi.org/10.1016/j.tmaid.2021.102044.
5. Alizon S, Haim-Boukobza S, Foulongne V, Verdurme L, Trombert-Paolantoni S, Lecorche E. Rapid spread of the SARS-CoV-2 Delta variant in some French regions. *Euro Surveill* 2021;**26**(28). pii=2100573. https://doi.org/10.2807/1560-7917.ES.2021.26.28.2100573.
6. Klatt M. What has happened to our cross-border regions? Corona, Unfamiliarity and transnational borderlander activism in the Danish-German border region. In: *Borders in Perspective*, vol. 4; 2020. p. 43–7. https://doi.org/10.25353/ubtr-xxxx-b825-a20b.
7. Duvernet C. Die Verbreitung von COVID-19 in Grenzregionen. *Inf Raumentwickl* 2021;**2**:50–9. Available at: https://elibrary.steiner-verlag.de/journal/izr/48/2.
8. Hossain MP, Junus A, Zhu X, Jia P, Wen TH, Pfeiffer D, et al. The effects of border control and quarantine measures on the spread of COVID-19. *Epidemics* 2020;**32**:100397. https://doi.org/10.1016/j.epidem.2020.100397.
9. Scarpone C, Brinkmann ST, Große T, Sonnenwald D, Fuchs M, Walker BB. A multimethod approach for county-scale geospatial analysis of emerging infectious diseases: a cross-sectional case study of COVID-19 incidence in Germany. *Int J Health Geogr* 2020;**19**:32. https://doi.org/10.1186/s12942-020-00225-1.
10. European Commission (EC). *Data from: member States' notifications of the temporary reintroduction of border control at internal borders pursuant to Article 25 and 28 et seq. of the Schengen Borders Code*. EC; June 13,2021. https://ec.europa.eu/home-affairs/what-we-do/policies/borders-and-visas/schengen/reintroduction-border-control_en.

Public Health 202 (2022) 1–9

Contents lists available at ScienceDirect

Public Health

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

Original Research

COVID-19 vaccination acceptability in the UK at the start of the vaccination programme: a nationally representative cross-sectional survey (CoVAccS – wave 2)

S.M. Sherman [a, *], J. Sim [b], M. Cutts [a], H. Dasch [c, d], R. Amlôt [e, f], G.J. Rubin [d, f], N. Sevdalis [c, d], L.E. Smith [d, f]

[a] Keele University, School of Psychology, UK
[b] Keele University, School of Medicine, UK
[c] King's College London, Centre for Implementation Science, UK
[d] King's College London, Institute of Psychiatry, Psychology and Neuroscience, UK
[e] Public Health England, Behavioural Science Team, Emergency Response Department Science and Technology, UK
[f] NIHR Health Protection Research Unit in Emergency Preparedness and Response, UK

ARTICLE INFO

Article history:
Received 20 May 2021
Received in revised form
6 October 2021
Accepted 9 October 2021
Available online 18 October 2021

Keywords:
Hesitancy
Side effects
Beliefs
Attitudes
Barriers
Covid-19 vaccines

ABSTRACT

Objectives: Investigate factors associated with the intention to have the COVID-19 vaccination following initiation of the UK national vaccination programme.
Study design: An online cross-sectional survey completed by 1500 adults (13th–15th January 2021).
Methods: Linear regression analyses were used to investigate associations between intention to be vaccinated for COVID-19 and sociodemographic factors, previous influenza vaccination, attitudes and beliefs about COVID-19 and COVID-19 vaccination and vaccination in general. Participants' main reasons for likely vaccination (non-)uptake were also solicited.
Results: 73.5% of participants (95% CI 71.2%, 75.7%) reported being likely to be vaccinated against COVID-19, 17.3% (95% CI 15.4%, 19.3%) were unsure, and 9.3% (95% CI 7.9%, 10.8%) reported being unlikely to be vaccinated. The full regression model explained 69.8% of the variance in intention. Intention was associated with: having been/intending to be vaccinated for influenza last winter/this winter; stronger beliefs about social acceptability of a COVID-19 vaccine; the perceived need for vaccination; adequacy of information about the vaccine; and weaker beliefs that the vaccine is unsafe. Beliefs that only those at serious risk of illness should be vaccinated and that the vaccines are just a means for manufacturers to make money were negatively associated with vaccination intention.
Conclusions: Most participants reported being likely to get the COVID-19 vaccination. COVID-19 vaccination attitudes and beliefs are a crucial factor underpinning vaccine intention. Continued engagement with the public with a focus on the importance and safety of vaccination is recommended.

Introduction

One year on from the emergence of COVID-19 in China in December 2019, there have been more than 112 million cases of COVID-19 and nearly 2.5 million deaths worldwide.[1] While countries have implemented a variety of public health measures to try to prevent the spread of the virus, scientists across the world have worked on developing effective vaccines. On 2nd December 2020, the United Kingdom (UK) became the first country to approve a COVID-19 vaccine that had been through a large-scale trial,[2] and on 8th December, the first dose of the Pfizer/BioNTech vaccine was administered.[3] This was swiftly followed by UK approval of the Oxford/AstraZeneca vaccine on 30th December 2020 and the Moderna vaccine on 8th January 2021. Given the severity of the pandemic and associated clinical outcomes, it is imperative that COVID-19 vaccination uptake is maximised so that, alongside ongoing protective public health practices, the spread of infection can be reduced.[4] To achieve this, we need to understand the factors that affect people's willingness to have a vaccine.

* Corresponding author. School of Psychology, Dorothy Hodgkin Building, Keele University, Keele, ST5 5BG, UK.
E-mail address: s.m.sherman@keele.ac.uk (S.M. Sherman).

https://doi.org/10.1016/j.puhe.2021.10.008
0033-3506/

The existing peer-reviewed research exploring the acceptability of a COVID-19 vaccination was all conducted before a vaccine was available,[5–8] when details about the actual vaccine were still a matter of speculation. For example, in a survey of 1500 UK adults that we conducted in July 2020,[5] 64% of participants reported being very likely to be vaccinated against COVID-19, 27% were unsure, and 9% reported being very unlikely to be vaccinated. Intention to be vaccinated was associated with: more positive general COVID-19 vaccination beliefs and attitudes; weaker beliefs that the vaccination would cause side effects or be unsafe; greater perceived information sufficiency to make an informed decision about COVID-19 vaccination; greater perceived risk of COVID-19 to others; older age; and having been vaccinated for influenza the previous year. Studies conducted before a vaccine was available provided useful data with which to start planning communication strategies about vaccine rollout. With national vaccination programmes currently underway internationally, further research is needed to understand how COVID-19 vaccine acceptance and factors affecting acceptance might have changed now that vaccination has materialised.

Contextual factors such as news stories and media coverage also influence vaccine acceptance.[9] The approval of the COVID-19 vaccines and the rollout of the vaccination programme has been accompanied by considerable press reporting of the differences between the vaccines, including the type of technology used (mRNA vs viral vector[10]), speculation about levels of efficacy observed in clinical trials, and potential variation of effectiveness in a public health context.[11] There was coverage related to two doctors in the United Kingdom who had an allergic reaction to the vaccine[12] and some controversy over the deviation from prior clinical trial administration of the required 2 doses of each vaccine 3 weeks apart so that they were administered 12 weeks apart,[13] as well as the potential to mix vaccine types.[14] These issues may also have influenced COVID-19 vaccine acceptance.

The aim of this study was to investigate associations between COVID-19 vaccination intention and sociodemographic, psychological, and contextual factors in a demographically representative sample of the UK adult population at the start of the COVID-19 vaccination programme rollout.

Methods

Design

We conducted an online cross-sectional survey (13th to 15th January 2021) hosted on Qualtrics.

Participants

Participants ($n = 1500$) were recruited through Prolific's online research panel and were eligible for the study if they were aged eighteen years or over, lived in the United Kingdom, and had not completed our previous survey[5] ($n > 31{,}000$ eligible participants). Prolific set quotas based on UK census data to ensure respondents were broadly representative of the UK population in terms of age, sex and ethnicity. Of 1508 people who began the survey, 1503 completed it (99.7% completion rate). Three participants were excluded from the sample as they did not meet quality control checks (specifically, they failed to correctly answer two or more of three attention check questions). Participants were paid £2 for a completed survey.

Measures

Full survey materials are available online.[15] Most items were the same as those in the UK survey reported above,[5] which was conducted in July 2020 and consisted of items that were based on previous literature.[16–20] Some further items were added, and some removed or amended to reflect the availability of specific COVID vaccinations and the timing of the survey.

Personal and clinical characteristics

We asked participants to report their age, gender, ethnicity, religion, highest educational or professional qualifications, current working situation, and total household income. We also asked participants what UK region they lived in, how many people lived in their household, whether they or someone else in their household (if applicable) had a long-standing illness, disability or infirmity and, if so, whether they had received a letter from the NHS recommending that they took extra precautions against coronavirus ('shielding') or whether they had a chronic illness that made them clinically vulnerable to serious illness from COVID-19. We also asked whether they or anyone they lived with were classified as obese or were pregnant and if they worked or volunteered in roles considered critical to the COVID-19 response ('key worker' roles).

Last, we asked participants whether they had been vaccinated for seasonal influenza last winter and/or had (or intended to be) this winter (yes/no).

Psychological and contextual factors

Participants were asked to what extent they thought 'coronavirus poses a risk to' people in the United Kingdom and to themselves personally, on a five-point scale, from 'no risk at all' to 'major risk'. They were asked if they thought they 'have had, or currently have, coronavirus'. Participants could answer 'I have definitely had it or definitely have it now', 'I have probably had it or probably have it now', 'I have probably not had it and probably don't have it now', and 'I have definitely not had it and definitely don't have it now'. They were also asked if they personally knew anyone who had had COVID-19 (yes/no).

Further, we asked participants a series of eight questions about their attitudes towards COVID-19. They were asked whether, as far as they knew, they were in one of the groups that had so far been offered the vaccine. Participants were then asked if they had been vaccinated (yes, I've had one/two doses/no), and if they answered yes, they were asked which vaccine they had received (Pfizer-BioNTech/Oxford University-AstraZeneca). All participants were then asked 21 questions about COVID-19 vaccination. Statements measured theoretical constructs, including perceived susceptibility to COVID-19, the severity of COVID-19, benefits of a COVID-19 vaccine, barriers to being vaccinated against COVID-19, ability to be vaccinated (self-efficacy), subjective norms, behavioural control, anticipated regret, knowledge, trust in the Government, and trust in the NHS. These items also investigated concerns about commercial profiteering and participants' beliefs about vaccination allowing life to get back to 'normal' and having to follow social distancing and other restrictions for COVID-19 if vaccinated. Participants rated the statements on an eleven-point scale (0–10) from 'strongly disagree' to 'strongly agree'. We adjusted the wording to make the grammatical tense either retrospective for those who had received the vaccine or prospective for those who had not. Participants who had not yet received a vaccine were additionally asked how likely they thought it was that they would get side effects from a coronavirus vaccine. We also asked participants if the coronavirus vaccination had been recommended to them by a health care professional and whether their employer did/would want them to have the COVID-19 vaccination. The order of items was quasi-randomised.

Outcome measure

To measure vaccination intention, we asked participants who had not yet been vaccinated to state how likely they would be to have a COVID-19 vaccination 'now that a coronavirus vaccination is available' on an eleven-point scale from 'extremely unlikely' (0) to 'extremely likely' (10).

We additionally asked participants to report the main reason why they were likely or unlikely to have a coronavirus vaccination in an open-text comment box.

Ethics

Ethical approval for this study was granted by Keele University's Research Ethics Committee (reference: PS-200129).

Sample size

A target sample size of 1500 was chosen to provide a high ratio of cases to estimated parameters in order to avoid overfitting and loss of generalisability in the regression model.[21]

Analysis

To identify variables associated with an intention to have the COVID-19 vaccination in those who had not yet been vaccinated, we constructed a linear regression model. Ordinal and multinomial predictors were converted to dummy variables. To aid interpretation of the model and to achieve a more parsimonious set of predictor variables, we ran principal component analyses[22] on items investigating beliefs and attitudes about a) COVID-19 and b) COVID-19 vaccination. This resulted in a smaller number of new variables (components) that are linear combinations of the original items and represent different latent dimensions that underlie these items. The components were generated separately for the items relating to COVID-19 as an illness and those relating to COVID-19 vaccination, and the components were named in accordance with the items that loaded most heavily upon them.

Variables entered into the regression model were selected *a priori* based on their theoretical relevance; no variable selection procedures were employed. Five groups of variables were included in the model: personal and clinical characteristics; seasonal influenza vaccination; general beliefs and attitudes relating to vaccination; beliefs and attitudes relating to COVID-19 illness; and beliefs and attitudes relating to COVID-19 vaccination. The percentage of variance in the outcome variable explained by each predictor was calculated as the squared semipartial correlation for a numerical or binary predictor and the change in R^2 attributable to a set of dummy variables.

As well as fitting the full model, we also added the groups of variables as successive blocks in a hierarchical model to determine the incremental increase in the adjusted R^2 value as these groups of variables were added to the model.

Due to the large number of predictors in the model, statistical significance was set at $P \leq .01$ to control Type 1 errors, and 99% confidence intervals (CIs) were correspondingly calculated for the regression coefficients. Assumptions of the analysis were checked. Analyses were conducted in SPSS 26.

To analyse open-ended responses for reasons why participants were likely or unlikely to have a coronavirus vaccination, we conducted a content analysis using an emergent coding approach, whereby codes were identified from the data rather than *a priori*.[23] Two authors (MC and HD) jointly coded a small sample of statements to understand the scope of the data. They then each independently coded sufficient responses that they achieved a run of 15 statements without encountering any new codes. At this point, they compared the codes they had generated and discussed any discrepancies. They then independently applied these codes to the rest of the statements, after which they checked that they had applied the same codes across the statements and discussed and resolved any additional codes and any discrepancies. This process was first applied to those participants who were uncertain about whether they would have the vaccine, then to those who were unlikely to have it, and finally to those participants who were likely to have it.

Results

Participants were broadly representative of the UK population (mean age 45.6 years, SD = 15.6, range 18–86; 51% female; 85% white ethnicity; Table 1, see Supplementary Materials 1 for further breakdown). At the time of completing the survey, only 30 respondents had received one or both doses of a coronavirus vaccine.

Descriptive statistics for items assessing psychological factors are reported in Tables 2 and 3. These tables show that participants were worried about catching coronavirus and did not believe that it would be a mild illness for them. Approximately three-quarters of participants (76.7%) believed COVID-19 posed a moderate or higher risk to them personally. It was also noteworthy that participants reported considerably more trust in the NHS compared to the Government regarding managing the pandemic.

Principal component analyses

Four components emerged from the principal component analysis on beliefs and attitudes about COVID-19, accounting for 75% of the variance in original items, and five components emerged from the principal component analysis investigating items related to a COVID-19 vaccination, accounting for 68% of the variance in the original items (see Supplementary Materials 2).

Vaccination intention

Participants' vaccination intention (in participants who had not already received one or both doses) is presented in Fig. 1. Vaccination intention exhibited a marked negative skew (mean = 8.13, standard deviation = 2.96, median = 10.00). In order to categorise respondents in terms of their vaccination intention, we applied *a priori* cut-points to the 0–10 scale (with scores of zero to two as 'very unlikely', three to seven as 'uncertain' and eight to ten as 'very likely', as per our July 2020 survey [5]). On this basis, 9.3% (95% CI 7.9%, 10.8%) reported being very unlikely to be vaccinated ($n = 136$), 17.3% (95% CI 15.4%, 19.3%) reported being uncertain about their likelihood of vaccination ($n = 254$), and 73.5% (95% CI 71.2%, 75.7%) reported being very likely to be vaccinated ($n = 1080$).

The final model explained 69.8% of the variance in intention to vaccinate (Table 4). Increased likelihood of being vaccinated for COVID-19 was significantly associated with having been vaccinated for influenza last or this winter (or intending to do so this winter), and with all of the components derived from the items relating to COVID-19 vaccination, other than '*freedom from restrictions through the vaccine*'. Vaccination intention also showed a significant negative association with beliefs that only people who are at risk of serious illness from coronavirus need to be vaccinated and that widespread coronavirus vaccination is just a way to make money for vaccine manufacturers.

The principal component that related to the necessity of vaccination explained more variance in vaccination intention than any other predictor in the statistical model, followed by the principal components concerning social acceptance and safety of the vaccine.

Table 1
Participant characteristics.

Personal and clinical characteristics	Level	*n* (%)
Sex	Male	728 (48.5)
	Female	765 (51.0)
	Other	6 (.1)
	Prefer not to say	1 (.1)
Ethnicity	White	1269 (84.6)
	Black and minority ethnic	224 (14.9)
	Prefer not to say	7 (.5)
Religion	No religion	793 (52.9)
	Christian	571 (38.1)
	Other religion	114 (7.5)
	Prefer not to say	22 (1.5)
Highest qualification	Degree equivalent or higher[c]	817 (54.5)
	Other or no qualifications	677 (45.1)
	Prefer not to say	6 (.4)
Employment status	Full-time	649 (43.3)
	Part-time	269 (17.9)
	Not working/other	572 (38.1)
	Don't know	1 (.1)
	Prefer not to say	9 (.6)
Key worker	Yes	500 (33.3)
	No	1000 (66.7)
Total household income[a]	Under £10,000	94 (6.3)
	£10,000–£19,999	215 (14.3)
	£20,000–£29,999	249 (16.6)
	£30,000–£39,999[b]	236 (15.7)
	£40,000–£49,999	179 (11.9)
	£50,000–£74,999	261 (17.4)
	£75,000 or over	161 (10.7)
	Don't know	18 (1.2)
	Prefer not to say	87 (5.8)
Region where respondent lives[a]	East Midlands	127 (8.5)
	East of England	111 (7.4)
	London	205 (13.7)
	North East	61 (4.1)
	North West	176 (11.7)
	Northern Ireland	27 (1.8)
	Scotland	116 (7.7)
	South East	239 (15.9)
	South West	131 (8.7)
	Wales	56 (3.7)
	West Midlands	122 (8.1)
	Yorkshire and the Humber	127 (8.5)
	Prefer not to say	2 (.1)
Number of people in household[a]	1	233 (15.5)
	2[b]	587 (39.1)
	3–4	563 (37.5)
	5–6	105 (7.0)
	7 or more	9 (.6)
	Prefer not to say	3 (.2)
Extremely clinically vulnerable – respondent	Yes	344 (22.9)
	No	1156 (77.1)
Extremely clinically vulnerable – other(s) in household	Yes	254 (16.9)
	No	1010 (67.4)
	Not applicable/prefer not to say	236 (15.7)
Influenza vaccination last winter[d]	Yes	457 (30.5)
	No	1040 (69.3)
	Don't know	1 (.1)
	Prefer not to say	2 (.1)
Influenza vaccine this winter[d]	Yes	581 (38.7)
	No, but intend to	180 (12.0)
	No, and don't intend to	723 (48.2)
	Don't know	13 (.9)
	Prefer not to say	3 (.2)

[a] Not included in regression model.
[b] Median category.
[c] Undergraduate (e.g. BA, BSc) or postgraduate (e.g. MA, MSc, PhD) degree or other technical, professional or higher qualification.
[d] Combined into a single variable in the regression model.

Other significant predictors only explained small percentages of variance.

When the groups of variables were entered hierarchically as blocks, we could infer the percentage of additional variance explained by each block from the change in incremental adjusted R^2. Personal and clinical characteristics (block 1) alone explained very little (8.8%) of the variance in intention to be vaccinated. When previous influenza vaccination (block 2) was added, it explained an additional 6.4% of the variance. Adding general vaccination beliefs and attitudes (block 3) resulted in the largest increase in percentage (25.1%) of explained variance (although in the full model, the predictors in this group were no longer significant). When beliefs and attitudes about COVID-19 (block 4) were added to the model, they explained 6.5% more of the variance in vaccination intention. Adding positive beliefs and attitudes about a COVID-19 vaccination (block 5) explained a further 23.0% of the variance. Each block explained a statistically significant percentage of the variance ($P < .001$ in each case).

Content analysis

Of the 1470 participants who had not yet received a vaccine and were asked to give a reason for the score provided on the likelihood of having the vaccination question, 1461 participants (99.4%) provided a response. Answers ranged from one word to 455 words (mean = 20.3, SD = 20.6). The content analysis generated 102 unique codes. The codes were then further organised into themes, and these, along with a frequency count of comments per theme, are presented in Table 5. A breakdown of codes and themes is provided in Supplementary Materials 3.

The two most frequently cited reasons to support the score participants gave on the likelihood question related to protecting themselves or others. These were primarily the reasons given by participants who indicated they were likely to have the vaccine. In comparison, the most frequently provided reasons provided by participants who were uncertain or unlikely to have the vaccine were related to safety concerns about the COVID-19 vaccine.

Discussion

The UK government has set a target of offering the first dose of a COVID-19 vaccine to all adults in the United Kingdom by the end of July 2021.[24] In this study conducted in January 2021, three-quarters of participants reported being very likely to have the vaccination. This is higher than the 64% who reported being very likely to have the vaccination in our study conducted in July 2020[5] and is consistent with increases in vaccination intention reported elsewhere. For example, in a recent (March 2021) YouGov poll in the UK, 86% of respondents had either already been vaccinated or reported that they would get the vaccine.[25] Despite the relatively high intention reported in our study and recent polling, we cannot be complacent about uptake. News stories are emerging almost every week about different variants of the virus, as well as issues around differential uptake of individual vaccines across the world,[26] and it is important to understand the factors associated with intention in order to maximise uptake and offset any adverse media reporting, social media misinformation, and the like.

Our results indicate that greater intention to have the COVID-19 vaccination was associated with having been vaccinated for influenza last or this winter or intending to be this winter. This is consistent with our previous findings,[5] as well as with findings from the US[27] and Europe.[28] However, several of the COVID-19

Table 2
Descriptive statistics for continuous items measuring beliefs and attitudes about COVID-19 and a COVID-19 vaccination and vaccination intention. Data are mean (standard deviation) on a 0–10 numerical rating scale (0 = strongly disagree, 10 = strongly agree), except where indicated. Also shown is the principal component (as numbered in Table 4) on which the items loaded most, for those items included in the principal components analysis (see Supplementary Materials 2).

	Item	Mean (SD)	Related principal component[b]
Attitudes and beliefs about COVID-19	I am worried about catching coronavirus	6.51 (2.82)	2
	I believe that coronavirus would be a mild illness for me	4.44 (2.73)	2
	Too much fuss is being made about the risk of coronavirus[a]	2.01 (2.63)	1
	We are all responsible for reducing the spread of coronavirus[a]	9.23 (1.59)	1
	I believe I am immune to coronavirus[a]	1.06 (1.91)	1
	The coronavirus pandemic has had a big impact on my life	7.51 (2.33)	4
	I trust the NHS to manage the coronavirus pandemic in the UK	7.39 (2.21)	3
	I trust the Government to manage the coronavirus pandemic in the UK	3.99 (2.98)	3
Attitudes and beliefs about a COVID-19 vaccination	How likely do you think it is that you would get side effects from a coronavirus vaccination (0 = very unlikely, 10 = very likely)	4.01 (2.45)	–
	A coronavirus vaccination should be mandatory for everyone who is able to have it	6.27 (3.60)	6
	Without a coronavirus vaccination, I am likely to catch coronavirus	6.45 (2.44)	6
	Two doses of coronavirus vaccination will protect me against coronavirus	7.52 (2.22)	1
	If I don't get the coronavirus vaccination and end up getting coronavirus, I will regret not getting the vaccination[a]	7.92 (3.04)	6
	It would be very easy for me to have a coronavirus vaccination[a]	7.81 (2.51)	1
	The coronavirus vaccination could give me coronavirus	1.59 (2.17)	7
	The way the coronavirus vaccines are being given in the UK goes against the manufacturers' recommendations	4.89 (2.99)	–
	I might regret getting the coronavirus vaccination if I later experienced side effects from it	4.42 (3.18)	7
	The coronavirus vaccination is too new for me to be confident about getting vaccinated	4.05 (3.28)	7
	Most people will get a coronavirus vaccination	7.46 (1.75)	5
	Other people like me will get a coronavirus vaccination[a]	7.94 (2.20)	5
	In general, vaccination is a good thing[a]	9.02 (1.72)	–
	I am afraid of needles[a]	2.77 (3.31)	–
	If I were vaccinated, I think I would not need to follow social distancing and other restrictions for coronavirus	2.60 (2.82)	9
	I know enough about the coronavirus *illness* to make an informed decision about whether or not to get vaccinated	7.73 (2.45)	8
	I know enough about the coronavirus *vaccine* to make an informed decision about whether or not to get vaccinated	6.79 (2.67)	8
	Only people who are at risk of serious illness from coronavirus need to be vaccinated	2.39 (3.04)	–
	My family would approve of my having a coronavirus vaccination[a]	8.58 (2.16)	5
	My friends would approve of my having a coronavirus vaccination[a]	8.33 (2.09)	5
	Widespread coronavirus vaccination is just a way to make money for vaccine manufacturers[a]	2.05 (2.62)	–
	The coronavirus vaccine will allow us to get back to 'normal'	7.24 (2.32)	5
Vaccination intention	Now that a coronavirus vaccination is available, how likely is it you will have one? (0 = very unlikely, 10 = very likely)[a]	8.13 (2.96)	–

[a] Skewed variables; mean values should be interpreted cautiously.
[b] 1 = perceived severity of COVID-19; 2 = individual vulnerability to COVID-19; 3 = trust in COVID-19 management; 4 = impact of COVID-19 on one's life; 5 = social norms; 6 = the necessity of vaccination; 7 = safety of the vaccine; 8 = adequacy of information about the vaccine; 9 = freedom from restrictions through the vaccine.

vaccination beliefs and attitudes explained a more substantial proportion of the variance in vaccination intention. Intention was associated with greater perceived social norms around COVID-19 vaccination and the greater perceived necessity of vaccination. These items map onto theoretical constructs that have previously been shown to influence the adoption of health behaviours: subjective norms and perceived susceptibility.[5] Previous studies exploring vaccine intentions have also found high levels of positive social norms favouring the vaccine[29] and that intention is associated with increased levels of concern related to the risks of the disease.[28,30] We also found that lower intention was associated with reduced belief in the safety of the vaccine, and this has been found consistently across studies exploring COVID-19 vaccine intentions[7,8] and vaccine hesitancy.[8] This was also reflected in the content analysis of participants' open-ended responses, in which issues related to vaccine safety were the most frequently identified reason for lower vaccination intention in the participants we classified as uncertain or very unlikely to have the vaccine. This is also consistent with the free-text responses given in an English study exploring vaccine acceptability conducted April to May 2020.[7] Since there was less than a year between the genetic code of COVID-19 being made public and the first COVID-19 vaccine being approved, this belief is perhaps unsurprising. However, it does make it essential that there is sufficient engagement with the public's concerns and that good-quality and credible information continues to be made available about the vaccine. This recommendation is reinforced by the association in our data of perceived adequacy of information about the vaccine to facilitate an informed decision with vaccination intention. Since free-text responses related to safety were the most frequent category of response from uncertain responders, it is likely that any reporting of safety concerns in the media may well shift the balance in favour of not being vaccinated, as has been observed previously.[31] The WHO guideline for emergency risk communication makes the strong recommendation that 'Communication by authorities to the public should include explicit information about uncertainties associated with risks, events and interventions, and indicate what is known and not known at a given time.'[32] However, it is imperative that halting the rollout of the vaccination programme because of unproven risks, as seen in some European countries in March 2021,[26] should be avoided, as this is likely to damage uptake once vaccination is restarted.

Vaccination intention in our study was also associated with a weaker belief that only people who are at risk of serious illness from coronavirus need to be vaccinated, and free-text responses associated with protecting others were frequently given to explain

Table 3
Descriptive statistics for categorical and ordinal items measuring beliefs, attitudes and behaviour relating to COVID-19 and a COVID-19 vaccination.

Item	Level	*n* (%)
To what extent do you think coronavirus poses a risk to people in the UK?	No risk at all	6 (.4)
	Minor risk	69 (4.6)
	Moderate risk	197 (13.1)
	Significant risk	568 (37.9)
	Major risk	659 (43.9)
	Don't know	1 (.1)
To what extent do you think coronavirus poses a risk to you personally?	No risk at all	37 (2.5)
	Minor risk	306 (20.4)
	Moderate risk	539 (35.9)
	Significant risk	424 (28.3)
	Major risk	187 (12.5)
	Don't know	7 (.5)
Do you believe you have had, or currently have, coronavirus?	Definitely not	460 (30.7)
	Probably not	621 (41.4)
	Probably	164 (10.9)
	Definitely	79 (5.3)
	Don't know	175 (11.7)
	Prefer not to say	1 (.1)
Do you personally know anyone (excluding yourself) who has had coronavirus?	Yes	1153 (76.9)
	No	336 (22.4)
	Don't know	10 (.7)
	Prefer not to say	1 (.1)
As far as you know, would your employer want you to have the coronavirus vaccination?	Yes	572 (60.6)
	No	59 (6.3)
	Don't know	313 (33.2)
	Not applicable	556

intention to have the vaccine. Statements about protecting others were only second in the content analysis to those relating to self-protection, such as a perceived high personal risk of disease severity, and this is consistent with previous research on COVID-19 vaccine intentions.[7] Since those individuals who are less likely to vaccinate believe that only those at risk need to be vaccinated, and those who are more likely to vaccinate favour protecting others, as well as themselves, it may be that messaging needs to be tailored to accommodate the possibility that emphasising the community benefits of vaccination may not motivate hesitant individuals.

Several previous studies have found that various sociodemographic factors are associated with COVID-19 vaccine intention, such as age,[5,30] gender[28,33] and ethnicity.[7,33] We did not find this in our study, and it is not entirely clear why that might be. The lack of impact of ethnicity on intention is perhaps the most striking absence, given both previous research and evidence from actual uptake in the United Kingdom, which shows that a significantly smaller proportion of ethnic minorities compared to white health care workers have had the COVID-19 vaccine.[34] We recruited a demographically representative sample based on age, gender and ethnicity. However, the relatively low number of participants from ethnic minority backgrounds necessitated collapsing our data across these categories, which may explain why ethnicity was not associated with intention in our study. The lack of an association with age is also unexpected, especially since by 19th September 2021, 73.9% of 18–24-year-olds had had at least one dose of the vaccine compared to 98.3% of 75–79-year-olds.[35] The age difference in actual uptake may in part be due to the way the vaccine was rolled out to the population, with the vaccine being distributed first to older age groups (25–29-year-olds first invited 6 months after distribution to residents in care homes and those aged 80 years and over).[36] The percentage of participants in our sample who indicated they were very likely to have the vaccine (73.5%) was lower than the percentage of the UK population who actually did (89.0% had had one dose and 82.9% had two doses by 19th September 2021[35]). This may reflect the fact that our study was conducted when the vaccine was first rolled out and was not yet available to most of our participants. As the rollout gathered pace, subjective norms and perceived susceptibility, both of which are shown to be important determinants of vaccine willingness in our study, may well have increased for older adults, with high uptake fuelling subjective norms and the disparity in reported health outcomes for vaccinated versus unvaccinated adults fuelling perceived susceptibility. Furthermore, the UK government and NHS have been proactive in providing mass vaccination centres, invitations, reminders, and messaging. A recent study found that younger adults in the UK demonstrated more vaccine hesitancy and resistance than older adults,[37] which may have prevented the same inflation in uptake relative to the reported likelihood that we have seen in older adults. Studies using appropriate sampling techniques are required to capture and quantify uptake and associated attitudinal differences across different population cohorts. As the vaccination is made available to all those who want it, future research could also usefully focus on those who are uncertain or unlikely to be vaccinated.

Our regression model explained 70% of the variance in vaccination intention, similar to the figure of 76% in our previous study[5]; the remaining unexplained variance will represent predictors that we did not measure and random measurement error. Given that social science research in general,[38] and public opinion surveys in particular,[39] do not characteristically yield high R^2 values owing to the likely multiple and complex determinants of individuals' beliefs and attitudes, we believe that our model demonstrates good explanatory power.

Finally, the intention to have the COVID-19 vaccine was associated with a weaker belief that widespread coronavirus vaccination is just a way to make money for vaccine manufacturers. This is consistent with research from Hong Kong,[40] in which higher levels of trust in the vaccine manufacturer were associated with increased willingness to have the COVID-19 vaccine.

Limitations of this study include that we measured self-reported intention rather than actual uptake. Intention is usually higher than uptake; however, vaccine intention predicts vaccine uptake and so acts as a useful proxy in the early stages of vaccination rollout. The survey is cross-sectional, so we are unable to infer causality between attitudinal factors and intention. Although our sample was

Table 4
Results of the full linear regression model analysing associations with vaccination intention (adjusted $R^2 = .698$). Parameter estimates relate to the full model containing all predictors. The unstandardised regression coefficients represent the change in likelihood of vaccination for a one-unit increase in the predictor variable (or, for dummy variables, a shift from the reference category to the category concerned). The figures under '% variance explained' represent the percentage of variance in the outcome variable uniquely explained by the item (or set of dummy variables) concerned. The model was based on 1401 cases with complete data.

Predictor variable	Level	Standardised coefficient	Unstandardised coefficient	99% confidence interval	*P* value	% variance explained
Block 1 – personal and clinical characteristics						
Age	Years	.046	.008	−.001, .018	.023	.11
Sex (reference: female)	Male	−.024	−.135	−.364, .094	.129	.05
Ethnicity (reference: black and minority ethnic)	White	−.002	−.016	−.378, .346	.907	<.01
Religion (reference: none)					.935	<.01
	Christian	.000	.003	−.249, .255		
	Other	.006	.068	−.414, .549		
Qualifications (reference: other)	Degree equivalent or higher	.028	.162	−.070, .393	.066	.07
Employment status (reference: not working/other)					.336	.05
	Part-time	.014	.105	−.236, .446		
	Full-time	−.014	−.080	−.356, .197		
Key worker (reference: not key worker)	Key worker	−.003	−.018	−.280, .245	.863	<.01
Extremely clinically vulnerable – self (reference: no)	Yes	−.023	−.155	−.444, .134	.166	.04
Extremely clinically vulnerable – household member (reference: no)	Yes	−.022	−.173	−.474, .129	.140	.05
Block 2 – previous influenza vaccination						
Did you/will you have a vaccination for influenza last/this winter? (reference: no)	Yes	.047	.270	.012, .528	.007[a]	.14
Block 3 – general vaccination beliefs and attitudes						
Vaccination is generally good (0–10)	0–10 scale	.011	.018	−.079, .116	.627	.01
I am afraid of needles (0–10)	0–10 scale	.016	.014	−.021, .049	.295	.02
Block 4 – beliefs and attitudes about COVID-19						
Perceived risk of COVID-19 to people in the UK (reference: major)					.822	.02
	None/minor	−.012	−.166	−.897, .565		
	Moderate	.007	.063	−.380, .507		
	Significant	.002	.015	−.261, .290		
Perceived risk of COVID-19 to oneself (reference: major)					.021	.21
	None/minor	.059	.407	−.190, 1.005		
	Moderate	.082	.493	.036, .951		
	Significant	.067	.429	.034, .824		
Do you have/have you had COVID-19? (reference: probably/definitely)					.288	.08
	Probably not	.028	.163	−.180, .586		
	Definitely not	−.002	−.010	−.375, .355		
	Don't know	.018	.164	−.283, .610		
Do you know anybody who has had COVID-19? (reference: no)	Yes	.005	.032	−.246, .310	.764	.02
Component 1: perceived severity of COVID-19		−.028	−.083	−.257, .091	.219	.03
Component 2: individual vulnerability to COVID-19		−.045	−.130	−.308, .048	.060	.08
Component 3: trust in COVID-19 management		−.005	−.015	−.144, .114	.766	<.01
Component 4: impact of COVID-19 on one's life		−.011	−.031	−.156, .093	.513	.01
Block 5 – beliefs and attitudes about COVID-19 vaccination						
Component 5: social norms		.378	1.106	.954, 1.259	<.001[a]	7.55
Component 6: the necessity of vaccination		.460	1.329	1.172, 1.487	<.001[a]	10.20
Component 7: safety of the vaccine		.370	1.068	.923, 1.212	<.001[a]	7.83
Component 8: adequacy of information about the vaccine		.148	.431	.308, .554	<.001[a]	1.76
Component 9: freedom from restrictions through the vaccine		.015	.043	−.077, .164	.353	.02
The way the coronavirus vaccines are being given in the UK goes against the manufacturers' recommendations		.005	.005	−.035, .045	.751	<.01
Only people who are at risk of serious illness from coronavirus need to be vaccinated	0–10 scale	−.064	−.061	−.106, −.016	.001[a]	.26
Widespread coronavirus vaccination is just a way to make money for vaccine manufacturers	0–10 scale	−.060	−.067	−.128, −.006	.004[a]	.18

[a] $p \leq .01$.

broadly representative across the dimensions of age, sex, and ethnicity, representativeness was not specifically sought in the quota sampling for other dimensions such as location in the United Kingdom or socio-economic status.

To our knowledge, this is the first peer-reviewed study investigating the intention to receive a COVID-19 vaccination in a demographically representative sample of the UK population since the COVID-19 vaccination rollout began in December 2020. Three-quarters of our sample reported being very likely to have the COVID-19 vaccination. However, since vaccine uptake may well be lower than vaccine intention, it is important to understand the factors associated with intention and to ensure that communication and engagement strategies related to the vaccination are informed by those factors. Going forward, it is not yet known for how long COVID-19 vaccines confer immunity or how effective they will continue to be against emerging strains as the virus mutates and, consequently, whether booster vaccines may be required.[24] In order to ensure the success of the current vaccination rollout and

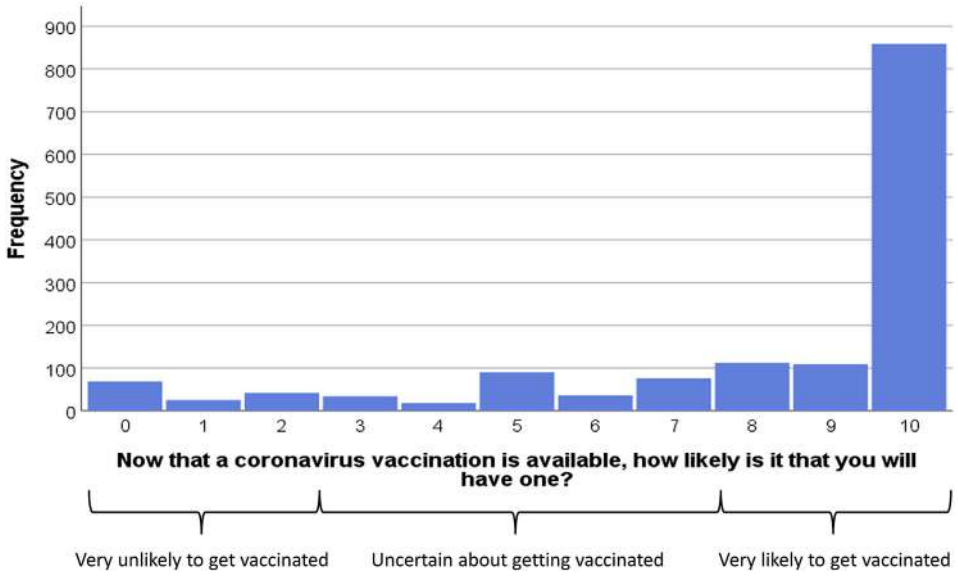


Fig. 1. Perceived likelihood of having a vaccination (0 = 'extremely unlikely' to 10 = 'extremely likely'). The figure also shows cut-points that we used to categorise respondents in terms of their vaccination intention.

Table 5
Thematic categorisation of codes generated by content analysis of reasons for likelihood of having, or not having, the COVID-19 vaccination, including a breakdown by likelihood of having the vaccination (likely, uncertain, likely). Themes are presented in descending order of overall frequency.

Theme name	Illustrative code	Number of comments per theme (likely, uncertain, unlikely)
Self-protection (including health reasons)	Perceived high personal risk of disease severity	675 (651, 22, 2)[a]
To protect others	Protecting the wider community	667 (609, 54, 4)[a]
To end the pandemic and its negative impacts	To end lockdown	345 (331, 14, 0)[a]
Confidence in vaccine and authority	The vaccine is effective	317 (276, 39, 2)[a]
Safety concerns about the COVID-19 vaccine	Concerns re quick development of the vaccine	226 (28, 110, 88)[b]
Concerns re details of vaccine (other than safety)	Concerns re dose time scale	173 (22, 89, 62)[b]
Low risk/no personal need for vaccine	Only high-risk groups need the vaccine	144 (25, 75, 44)[b]
Concern about health effects of COVID-19	Concerns re long term side effects of virus	77 (71, 6, 0)[a]
Other (miscellaneous)	Currently no access due to visa status	67 (46, 13, 8)[a]
Precontemplation/not made decision	Not offered yet	64 (12, 44, 8)[b]
Unspecified concerns about COVID-19 vaccine	Anxiety re the vaccine	62 (0, 24, 38)[b]
To travel/move around more freely	Wanting to visit family	55 (42, 13, 0)[a]
Specific health concerns about the COVID-19 vaccine	Fertility concerns	39 (5, 19, 15)[b]
Avoid/delay having the vaccine by waiting	Wanting others to test it first	36 (5, 25, 6)[b]
Lack of trust in authority	Lack of trust in media transparency	33 (7, 7, 19)[b]
General vaccine concerns	Fear of needles	30 (4, 8, 18)[b]

[a] Comments were made more frequently by those who were likely to vaccinate.
[b] Comments were made more frequently by those who were uncertain or unlikely to vaccinate.

any subsequent vaccination waves, our findings underline the importance of ongoing clear communication informed by theoretical constructs related to COVID-19 vaccination beliefs and attitudes, and the need for such communication to emphasise social acceptance of the vaccination, the importance of vaccination to stop the spread of COVID-19, even in the absence of underlying risk factors, and the safety of the vaccination.

Author statements

Ethical approval

Ethical approval for this study was granted by Keele University's Research Ethics Committee (reference: PS-200129).

Funding

The authors disclose receipt of the following financial support for the research, authorship, and/or publication of this article: a Keele University Faculty of Natural Sciences Research Development award to SS, JS and NS, and a Kings COVID Appeal Fund award granted jointly to LS, GJR, RA, NS, SS and JS. LS, RA and GJR are supported by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Emergency Preparedness and Response, a partnership between Public Health England, King's College London and the University of East Anglia. NS′ research is supported by the National Institute for Health Research (NIHR) Applied Research Collaboration (ARC) South London at King's College Hospital NHS Foundation Trust. NS is a member of King's Improvement Science, which offers co-funding to the NIHR ARC South London and is funded by King's Health Partners (Guy's and St Thomas' NHS Foundation Trust, King's College Hospital NHS Foundation Trust, King's College London and South London and Maudsley NHS Foundation Trust), and the Guy's and St Thomas' Charity. The views expressed are those of the authors and not necessarily those of the NIHR, the charities, Public Health England or the Department of Health and Social Care.

Competing interests

NS is the director of the London Safety and Training Solutions Ltd, which offers training in patient safety, implementation solutions and human factors to healthcare organizations and the pharmaceutical industry. The other authors have no conflicts of interest to declare.

Transparency declaration

The authors affirm that the manuscript is an honest, accurate, and transparent account of the study being reported and no important aspects of the study have been omitted; and that any discrepancies from the study as originally planned have been explained.

Data sharing statement

Data are available online[15].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2021.10.008.

References

1. https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases. [Accessed 13 March 2021].
2. Ledford H, Cyranoski D, Van Noorden R. The UK has approved a COVID vaccine — here's what scientists now want to know. *Nature* 2020;**588**:205–6. https://doi.org/10.1038/d41586-020-03441-8.
3. https://www.bbc.co.uk/news/uk-55227325. [Accessed 13 March 2021].
4. Axelsen PH, Poland GA. Vaccines, masks, distancing and credibility: an urgent warning for pandemic management. *Vaccine* 2021;**39**:1173–4. https://doi.org/10.1016/j.vaccine.2021.01.052.
5. Sherman SM, Smith L, Sim J, Amlôt R, Cutts M, Dasch H, et al. COVID -19 vaccination intention in the UK: results from the COVID-19 vaccination acceptability study, a nationally representative cross-sectional survey. *Hum Vaccin Immunother* 2021;**17**:1612–21. https://doi.org/10.1080/21645515.2020.1846397.
6. Freeman D, Loe BS, Chadwick A, Vaccari C, Waite F, Rosebrock L, et al. COVID-19 vaccine hesitancy in the UK: the Oxford coronavirus explanations, attitudes, and narratives survey (Oceans) II. *Psychol Med* 2020. https://doi.org/10.1017/S0033291720005188. online ahead of print.
7. Bell S, Clarke R, Mounier-Jack S, Walker JL, Paterson P. Parents' and guardians' views on the acceptability of a future COVID-19 vaccine: a multi-methods study in England. *Vaccine* 2020;**38**:7789–98. https://doi.org/10.1016/j.vaccine.2020.10.027.
8. Sallam M. COVID-19 vaccine hesitancy worldwide: a concise systematic review of vaccine acceptance rates. *Vaccines* 2021;**9**:160. https://doi.org/10.3390/vaccines9020160.
9. Larson HJ, Jarrett C, Eckersberger E, Smith DM, Paterson P. Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007–2012. *Vaccine* 2014;**32**:2150–9. https://doi.org/10.1016/j.vaccine.2014.01.081.
10. https://www.bbc.co.uk/news/health-55586410. [Accessed 13 March 2021].
11. https://www.ft.com/content/ac5e5ef8-bccb-482b-9f8d-0dab5cac6f9a. [Accessed 13 March 2021].
12. https://www.telegraph.co.uk/global-health/science-and-disease/coronavirus-news-vaccine-pfizer-nhs-oxford-covid-uk-cases/. [Accessed 13 March 2021].
13. https://www.theguardian.com/world/2021/jan/05/no-data-to-support-uk-delay-of-vaccines-second-dose-says-who. [Accessed 13 March 2021].
14. https://www.theguardian.com/society/2021/jan/03/analysis-is-it-wise-for-england-to-mix-and-match-covid-vaccines. [Accessed 13 March 2021].
15. https://osf.io/dy5jv/.
16. Rubin GJ, Bakhshi S, Amlôt R, Fear N, Potts HWW, Michie S. The design of a survey questionnaire to measure perceptions and behaviour during an influenza pandemic: the Flu TElephone Survey Template (FluTEST). *Health Serv Deliv Res* 2014;**2**:41. https://doi.org/10.3310/hsdr02410.
17. Lee Mortensen G, Adam M, Idtaleb L. Parental attitudes towards male human papillomavirus vaccination: a pan-European cross-sectional survey. *BMC Public Health* 2015;**15**:624. https://doi.org/10.1186/s12889-015-1863-6.
18. Myers LB, Goodwin R. Determinants of adults' intention to vaccinate against pandemic swine flu. *BMC Public Health* 2011;**11**:15. https://doi.org/10.1186/1471-2458-11-15.
19. Wheelock A, Miraldo M, Thomson A, Vincent C, Sevdalis N. Evaluating the importance of policy amenable factors in explaining influenza vaccination: a cross-sectional multinational study. *BMJ Open* 2017;**7**:e014668. https://doi.org/10.1136/bmjopen-2016-014668.
20. Smith LE, Mottershaw AL, Egan M, Waller J, Marteau TM, Rubin GJ. The impact of believing you have had COVID-19 on self-reported behaviour: cross-sectional survey. *PLoS One* 2020;**15**:e0240399. https://doi.org/10.1371/journal.pone.0240399.
21. Riley RD, Snell KIE, Ensor J, Burke DL, Harrell FE, Moons KGM, et al. Minimum sample size for developing a multivariable prediction model: part I — Continuous outcomes. *Stat Med* 2019;**38**:1262–75. https://doi.org/10.1002/sim.7993.
22. Bartholomew DJ, Steele F, Moustaki I, Galbraith JI. *The analysis and interpretation of multivariate data for social scientists*. Boca Raton, FL: Chapman & Hall/CRC; 2002.
23. Stemler S. An overview of content analysis. *Practical Assess Res Eval* 2000;**7**:17. https://doi.org/10.7275/z6fm-2e34.
24. Majeed A. *Can the UK meet its target of offering all adults a COVID-19 vaccine by 31 July?*. 2021. https://blogs.bmj.com/bmj/2021/02/22/can-the-uk-meet-the-governments-target-of-offering-all-adults-a-covid-19-vaccine-by-31-july/. [Accessed 24 March 2021].
25. *YouGov (12th March 2021)*. https://yougov.co.uk/topics/health/survey-results/daily/2021/03/12/66e85/3. [Accessed 22 March 2021].
26. https://www.theguardian.com/society/2021/mar/15/which-european-states-have-paused-astrazeneca-jabs-due-to-clotting-concerns. [Accessed 22 March 2021].
27. Ruiz JB, Bell RA. Predictors of intention to vaccinate against COVID-19: results of a nationwide survey. *Vaccine* 2021;**39**:1080–6. https://doi.org/10.1016/j.vaccine.2021.01.010.
28. Patelarou E, Galanis P, Mechili EA, Argyriadi A, Argyriadis A, Asimakopoulou E, et al. Factors influencing nursing students' intention to accept COVID-19 vaccination: a pooled analysis of seven countries. *Nurse Educ Today* 2021;**104**:105010. https://doi.org/10.1016/j.nedt.2021.105010.
29. Seale H, Heywood AE, Leask J, Sheel M, Durrheim DN, Bolsewicz K, et al. Examining Australian public perceptions and behaviors towards a future COVID-19 vaccine. *BMC Infect Dis* 2021;**21**:120. https://doi.org/10.1186/s12879-021-05833-1.
30. Lackner CL, Wang CH. Demographic, psychological, and experiential correlates of SARS-CoV-2 vaccination intentions in a sample of Canadian families. *Vaccine X* 2021;**8**:100091. https://doi.org/10.1016/j.jvacx.2021.10009.
31. Suppli CH, Hansen ND, Rasmussen M, Valentiner-Branth P, Krause TG, Mølbak K. Decline in HPV-vaccination uptake in Denmark — the association between HPV-related media coverage and HPV-vaccination. *BMC Public Health* 2018;**18**:1360. https://doi.org/10.1186/s12889-018-6268-x.
32. World Health Organization. *Communicating risk in public health emergencies: a WHO guideline for emergency risk communication (ERC) policy and practice*. Geneva: World Health Organization; 2017.
33. Robertson E, Reeve KS, Niedzwiedz CL, Moore J, Blake M, Green M, et al. Predictors of COVID-19 vaccine hesitancy in the UK household longitudinal study. *Brain Behav Immun* 2021;**94**:41–50. https://doi.org/10.1016/j.bbi.2021.03.008.
34. Martin CA, Marshall C, Patel P, Goss C, Jenkins DR, Ellwood C, et al. Association of demographic and occupational factors with SARS-CoV-2 vaccine uptake in a multi-ethnic UK healthcare workforce: a rapid real-world analysis. *medRxiv* 2021. https://doi.org/10.1101/2021.02.11.21251548.
35. NHS. *COVID-19 vaccinations*. 2021. https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-vaccinations/. [Accessed 28 September 2021].
36. NHS. *News: 25 to 29-year olds to be invited to receive life-saving Covid vaccine*. 2021. https://www.england.nhs.uk/midlands/2021/06/07/25-to-29-year-olds-to-be-invited-to-receive-life-saving-covid-vaccine/. [Accessed 4 October 2021].
37. Murphy J, Vallières F, Bentall RP, Shevlin M, McBride O, Hartman TK, et al. Psychological characteristics associated with COVID-19 vaccine hesitancy and resistance in Ireland and the United Kingdom. *Nat Commun* 2021;**12**:1–15.
38. Taraday M, Wieczorek-Taraday A. R-squared. In: Frey BB, editor. *The SAGE encyclopedia of educational research, measurement, and evaluation*, vol. 3; 2021. p. 1631–3.
39. Lewis-Beck MS. R-Squared. In: Lewis-Beck MS, Bryman A, Liao TF, editors. *The SAGE encyclopedia of social science research methods*, vol. 3; 2004. p. 983–4.
40. Wong MC, Wong EL, Huang J, Cheung AW, Law K, Chong MK, et al. Acceptance of the COVID-19 vaccine based on the health belief model: a population-based survey in Hong Kong. *Vaccine* 2021;**39**:1148–56. https://doi.org/10.1016/j.vaccine.2020.12.083.

Public Health 202 (2022) 10–11

Contents lists available at ScienceDirect

Public Health

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

Editorial

COVID-19 vaccine dilemmas

What a difference a year makes. We have seen the emergence of a novel zoonotic virus producing a global pandemic that has so far caused more than 92 million infections and two million deaths worldwide[1] with no signs of abating, despite a plethora of non-pharmacological measures deployed against it. But there is hope – the wonders of modern vaccine science have seen the rapid development of more than 68 vaccines worldwide, and around ten have received emergency authorisation and use thus far.[2]

This has opened a new front in the struggle to control the pandemic, offering the potential to achieve population immunity through vaccination. Vaccination is far safer than the more hazardous route of achieving population immunity through natural infection that carries a high risk of COVID-19 mortality and morbidity. Early experience with the vaccine is positive when compared directly with the effects of COVID-19 infection. For example, the roll-out of the Pfizer-BioNTech vaccine in the US observed only 21 cases of anaphylaxis after administration of nearly two million first doses of the vaccine, with no fatalities reported.[3] This compares favourably against the COVID-19 infection-to-fatality ratio, estimated at around 1.15% in high-income countries.[4] There is also a significant morbidity risk with COVID-19, including the risk of 'long COVID', that is as yet poorly understood. UK estimates are that around one in five persons infected with COVID-19 exhibit symptoms for a period of 5 weeks or longer, and one in ten respondents have symptoms for over 12 weeks.[5] From a population health perspective, there can be no rational reason for pursuing a population immunity strategy through natural infection now.

The advent of COVID-19 vaccines, however, has also created dilemmas. In the UK, faced with a rapidly spreading third wave of infections in December, partly driven by the emergence of a new variant (SARS-CoV-2 B.1.1.7), the Government switched from delivering the authorised two-dose schedule to prioritising first-dose coverage and delaying the second dose from 3–4 weeks to 12 weeks. This generated considerable uproar among primary care physicians involved in the delivery of vaccinations for a variety of reasons, including the turmoil and workload associated with having to consent patients and rebook hundreds of thousands of appointments.

The first dose vs two-dose prioritisation saga is also an ethical dilemma for clinicians. Clinicians typically strive to do their very best for individual patients and see it as their moral duty to do so. Giving two doses as per the vaccine authorisation and trial protocol could be seen as the 'right' thing to do. This approach has an absolutist lens as well and could be perceived as a choice between right vs wrong. Consequently, our natural tendency would be to follow the vaccine trial protocols, medical licensing and manufacturers' instructions as there is a 'certainty' to this. Failure to do so leads to an understandable concern that patients would be receiving suboptimal protection and substandard care that is not in line with best practice.

The counter perspective is the utilitarian view of the greatest good for the greatest number. A single dose would save more lives. Where resources are limited, there will be this trade-off. Prioritising two doses for some patients means denying others the protection that the first dose affords. Indeed, further analysis of the vaccine trial data suggests that the first dose of the Pfizer-BioNTech vaccine would afford patients around 89% protection 14 days after vaccination, and the second dose would only provide a marginal gain to 95%.[6]

The utilitarian approach tends to align with the population health approach as the perspective is of the welfare of groups of people rather than individuals. This conflicts with the patient-centric values that most clinicians have. Done right, the population health approach saves lives. The issue with this approach is we may not always know who we have saved, and those saved are unlikely to know they have been saved. It is easier to feel guilty for letting down the patient you have seen who has to be told their second dose has been delayed than the patient you have not yet seen whose first dose has been delayed. The two-dose vaccine prioritisation approach, with the limited number of vaccines, means only half the number of people getting vaccinated for the same number of available vaccines. If viral infections are spreading slowly, there is the luxury of time, and we can adopt the two-dose schedule for the most vulnerable and let other patient groups wait. However, faced with a worsening situation in the UK with a more transmissible virus, the only expedient option was to pursue a first dose prioritisation approach in the expectation that it would save more lives.

The other concern raised by those averse to the first dose approach was that this could lead to more vaccine failure or potentially introduce a selection pressure that favours mutant variants to emerge resistant to the vaccine, i.e. vaccine escape. Reassuringly, the view from immunology experts is that delaying the second dose by 8 weeks is unlikely to have a negative effect on the overall immune response. Neither is such an approach anticipated to lead to any specific safety issues to arise for the individual.[7] Indeed, it can also be argued that higher numbers of infection increase the likelihood of viral mutation, and consequently, efforts to reduce infection numbers may be more important for averting the risk of vaccine escape.

Another vaccine dilemma that has emerged is the decision as to who gets immunised first. The US and UK have both focused initially on the older age groups owing to their risk of mortality. One modelling study supports this approach and found that vaccine prioritisation for the elderly saves the most lives.[8] However, although vaccinating the elderly may reduce the number of deaths and hospital admissions, this age group accounts for only a small

https://doi.org/10.1016/j.puhe.2021.01.009

proportion of infections. Consequently, the impact on disease transmission in the community may be limited.

Indonesia, on the other hand, has adopted a different approach to mass COVID-19 vaccination, with a focus on working-age adults instead of the elderly in an attempt to revive its economy.[9] It is recognised that working-age adults generally mix more, and thus, this approach could decrease community transmission faster. In turn, this could provide a degree of protection to more vulnerable unvaccinated individuals. For now, it is unclear which approach will work best, and it will be interesting to compare the impacts of the different approaches on disease transmission and mortality in the coming months. It should also be remembered that a 'one-size-fits-all' approach rarely works as the social, political, economic and health system contexts will differ between countries. What is best for one country may not be best for another.

Finally, the arrival of COVID-19 vaccines has sparked a vaccine race between countries to immunise their populations in the hope it may restore some semblance of normality afterwards. This race favours high-income countries, and there are real concerns that vaccine nationalism could undermine cooperative efforts to control the pandemic globally.[10] This will create losers and widen global inequalities.

Mass vaccination in high-income countries does not necessarily confer security as there remains the risk of reimportation of infections from lower income countries where the virus is endemic.[11] There is also a moral dimension — is it right to vaccinate large numbers of predominantly lower risk individuals in high-income countries over other vulnerable individuals elsewhere? Indeed, should vaccine access not be determined by need rather than national wealth and influence? This is perhaps why the COVAX initiative is vital to ensuring equity of vaccine access.[12]

In an interconnected globalised world, all our fates are intertwined. Global solidarity is needed to protect our national health, wealth and human rights. In essence, we are not safe until we are all safe.

References

1. WHO. Coronavirus disease (COVID-19) dashboard [website]. https://covid19.who.int/Accessed 16/1/2021.
2. NY Times. Coronavirus vaccine tracker [webpage]. https://www.nytimes.com/interactive/2020/science/coronavirus-vaccine-tracker.html Accessed 16/1/2021.
3. Shimabukuro TT. Allergic reactions including anaphylaxis after receipt of the first dose of Pfizer-BioNTech COVID-19 vaccine—United States, December 14–23, 2020. *MMWR Morb Mortal Wkly Rep* 2021;**70**:46–51. https://doi.org/10.15585/mmwr.mm7002e1.
4. Brazeau N, Verity R, Jenks S, Fu H, Whittaker C, Winskill P, Dorigatti I, Walker P, Riley S, Schnekenberg RP, Heltgebaum H. Report 34: COVID-19 infection fatality ratio: estimates from seroprevalence.[online] https://spiral.imperial.ac.uk/handle/10044/1/83545 Accessed 16/1/2021.
5. Office for National Statistics. *The prevalence of long COVID symptoms and COVID-19 complications*. ONS; 16 Dec 2020 [online], https://www.ons.gov.uk/news/statementsandletters/theprevalenceoflongcovidsymptomsandcovid19complications. [Accessed 16 January 2021].
6. *Pfizer-BioNTech COVID-19 Vaccine (BNT162, PF-07302048) Vaccines and related biological products advisory committee: briefing document*. FDA; 10 Dec 2020 [online], https://www.fda.gov/media/144246/download. [Accessed 16 January 2021].
7. British Society for Immunology. *British Society for Immunology statement on COVID-19 vaccine dosing schedules*. 4 January 2021 [online], https://www.immunology.org/policy-and-public-affairs/briefings-and-position-statements/COVID-19-vaccine-dosing-schedules. [Accessed 16 January 2021].
8. Moore S, Hill EM, Dyson L, Tildesley M, Keeling MJ. *Modelling optimal vaccination strategy for SARS-CoV-2 in the UK*. medRxiv; 2020 Jan 1. https://www.medrxiv.org/content/10.1101/2020.09.22.20194183v2.full.pdf. [Accessed 16 January 2021].
9. Reuters. *Why Indonesia is vaccinating its working population first, not elderly*. 4 January 2021 [online], https://www.reuters.com/article/us-health-coronavirus-indonesia-explaine/why-indonesia-is-vaccinating-its-working-population-first-not-elderly-idUSKBN2990MX. [Accessed 16 January 2021].
10. Bollyky TJ, Bown CP. The tragedy of vaccine nationalism: only cooperation can end the pandemic. *Foreign Aff* 2020;**99**:96.
11. Financial Times. Letter: west's vaccine rollout must not ignore needs of poorer nations.[online] https://www.ft.com/content/bc9e2ad7-2bb1-4109-a868-c16e13735c3a Accessed 16/1/2021.
12. WHO. COVAX: working for global equitable access to COVID-19 vaccines.[website] https://www.who.int/initiatives/act-accelerator/covax Accessed 16/1/2021.

A.C.K. Lee*
The University of Sheffield, UK

J.R. Morling
The University of Nottingham, UK

* Corresponding author.
E-mail address: MDavis@rsph.org.uk (A.C.K. Lee).

Available online 1 February 2021

Public Health 202 (2022) 18–25

Contents lists available at ScienceDirect

Public Health

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

Original Research

Do children in India grow well into adolescents? Longitudinal analysis of growth transitions from Young Lives panel survey in India

S.K. Das [a], A.D. Burma [b], S. Amudhan [b, *], V. Mishra [c], P. Mahapatra [b], A. Ashok [b], M. Philip [a]

[a] Department of Biostatistics, National Institute of Mental Health and Neuro Sciences, Bangalore, India
[b] Department of Epidemiology, National Institute of Mental Health and Neuro Sciences, Bangalore, India
[c] Consultant-Rapid Response Team, WHO, India

ARTICLE INFO

Article history:
Received 18 June 2021
Received in revised form
6 October 2021
Accepted 19 October 2021
Available online 4 December 2021

Keywords:
Child malnutrition
Stunting
Adolescent
Longitudinal surveys
Markov chains

ABSTRACT

Objectives: Studies that examined the growth during late childhood and early adolescence beyond 8 years of age are very limited. Further, most studies have used dichotomized classification of stunting, thereby limiting the understanding of moderate stunting in childhood growth trajectory. We aimed to examine the course of stunting from childhood to adolescence by undertaking robust analyses of the Young Lives Survey (YLS) longitudinal data from India using multilevel categorization of stunting.

Study design: Retrospective cohort analysis was undertaken from YLS in India among 1827 children from the younger cohort born in 2001–02 with complete follow-up data in all five rounds of YLS collected in 2002, 2006, 2009, 2013, and 2016.

Methods: A three-state multistate Markov model (not stunted, moderate, severe) was performed to estimate annual transition probabilities, mean sojourn-time, and transition-specific risk factors.

Results: Between Round-one and Round-five, cross-sectional prevalence of severe stunting decreased from 10.4% (95% confidence interval [CI]: 7.8%, 13.7%) to 5.3% (95% CI: 3.8%, 7.3%), while moderate stunting increased from 19.9% (95% CI: 16.3%,23.9%) to 21.7% (95% CI: 18.4%, 24.9%). Mean Sojourn time estimation indicated a relatively concise state for moderate stunting. The stunting trajectory had shown gender differential where more faltering to severe stunting and lower recovery to the normal state was observed among girls between 8 and 12 years and among boys between 12 and 15 years. Compared with boys, girls had 40% excess likelihood (Hazard Ratio: 1.40; 95% CI 1.00 to 1.95) for moderate-to-severe stunting transition and also had 19% excess likelihood (Hazard Ratio: 1.19; 95% CI 1.01 to 1.40) of favorable transition (moderate-to-non-stunted).

Conclusions: The transition trajectory highlights preadolescence, especially among girls, as an additional window of opportunity to ensure better nutrition in adolescent life. With a fifth of adolescents living in India, study findings call for coordinated, multisectoral, age-appropriate, and gender-responsive approach to take India closer to meeting SDG-2.

Introduction

India has a huge and persistent burden of malnutrition in children.[1] Despite India's impressive economic growth, the prevalence of childhood stunting is still alarmingly high (38.4%) when compared to other developing countries.[2,3] Stunting, a sign of chronic undernutrition, is defined as height-for-age more than two standard deviations below (<-2SD) the WHO Child Growth Standards median.[3] At 46.6 million, India is home to the highest number of stunted children in the world, accounting for nearly one-third of the global burden of under-five childhood stunting.[2] Among the various recognized indicators of child nutritional status, stunting is used widely to monitor public health and nutrition program effectiveness and is prioritized as one of the six global nutrition targets for catalyzing global change. Early childhood stunting was associated with long-term adverse impacts on adolescents, making them more susceptible to disease risk, poor educational performance, and poverty in low-income countries as they grow into adults.[4–6] Thus, understanding the future course and determinants of childhood stunting in India becomes critical

* Corresponding author. 212, Department of Epidemiology, Dr MVG Building, NIMHANS, 560029, India. Tel.: +91 9972419985; Fax: +91 80 26564830.
E-mail address: sam_mmc1999@yahoo.co.in (S. Amudhan).

https://doi.org/10.1016/j.puhe.2021.10.010
0033-3506/

for UN Sustainable Development Goals (SDGs) that targets at eliminating malnutrition by 2030.

Previously, stunting during early childhood was often regarded as irreversible, and children with normal growth were assumed not at-risk of being stunted later in life.[7,8] This has led to excessive focus on the first 1000 days with shift in focus from interventions outside this period.[9] Most of these conclusions were based on cross-sectional data.[10] There are relatively few studies that examined childhood growth from a longitudinal perspective and notably only a few studies from India that examined the child growth up to 8 years.[10,11] These studies highlighted the potential for recovery and incident stunting after first 1000 days. However, there are only a few studies that examined the growth during late childhood and early adolescence beyond 8 years of age.[10,12]

It has been highlighted that moderate stunting is associated with a large proportion of nutrition-related complications than severe stunting.[13] Yet, most of the available studies from India have combined moderate and severe stunting into a single category.[10,14,15] Consequently, there is a lack of evidence on the temporal course of growth and the role of moderate stunting that influences late childhood and adolescence in India. Given the limitations of previous studies, we aimed to examine the course of stunting (temporal change in prevalence, annual transition probabilities, and determinants) from childhood to adolescence by undertaking robust analyses of Young Lives longitudinal data from India using multilevel categorization of stunting.

Methods

Data source

The data for the present retrospective cohort analysis were drawn from all five rounds of the Young Lives Survey (YLS) in India. YLS is an international panel study of children carried out over 15 years in four low-income and-middle-income countries, namely Ethiopia, Peru, Vietnam, and India.[16] In India, YLS was conducted in the states of Andhra Pradesh and Telangana during the years 2002, 2006–2007, 2009 and 2013–14 and 2016–17.

Details of the YLS sampling procedure and data collection are provided elsewhere.[16] Briefly, YLS in India was undertaken in a representative group of six districts from the former undivided state of Andhra Pradesh. By using sentinel site methodology, 20 sentinel sites/mandals were chosen from the selected districts to represent a mixture of poor, non-poor, rural and urban sites. Among the selected sentinel sites, households having a one-year-old child (born in 2001–02) or an eight-year-old child (born in 1994–95) were chosen randomly. The selection of households was made in such a way that, from each sentinel site there were 100 and 50 households with a child for a cohort of 2001–02 and 1994–95 respectively.[16] As the older cohort had a very high proportion of missing data in round 4, only the younger cohort born in 2001–02 was included for the current study (Supplementary Fig. S1).

Outcome measures

The outcome of interest considered for the present study is stunting. YLS measured supine length and height measurements to the nearest millimeter for children aged one year and aged 5,8,12, and 15 years, respectively, using standardized length and stadiometers. Height-for-age z (HAZ) scores were then calculated using the WHO Growth Standards (for children less than 5 years), and WHO Growth References (for children aged over 5 years).[17,18] Based on HAZ scores, stunting was classified into three mutually exclusive categories, namely not stunted (HAZ > −2), moderately stunted (HAZ between −3 and −2), and severely stunted (HAZ < −3). This multilevel categorization would enable a more nuanced assessment of the effect of different grades of stunting.

Quality control

The fieldworkers were thoroughly trained in anthropometric measurement and recording. The fieldworkers were monitored on a day-to-day basis by supervisors using a strict protocol.[16] Supervisors also observed a proportion of interviews to check the approach used by fieldworkers to ensure quality control. Data were checked for any inconsistencies.

Covariates

The covariates for childhood stunting considered for the present study included gender (Boy/Girl); place of residence (Rural/Urban); Religion (Hindu/Others); social group (Scheduled caste/Scheduled Tribe/Others); education of the mother (No formal Education/Upto grade 5/More than grade 5), household size (Upto 4/More than 4) and household wealth tertile (Poor/Middle/Rich). The household wealth tertile was constructed from the household wealth index available from the original data.[19] Briefly, wealth index includes a broad range of variables as markers of wealth and is computed as a simple average of the three indices: housing quality, access to services, and ownership of consumer durables. The value of the wealth index range between 0 and 1, where a higher wealth index indicates a higher socio-economic status.[19] The gender, religion, mother's education at birth, and social group were considered as time-independent covariates, while the place of residence, household size, and household wealth tertile were considered as time-dependent covariates.

Statistical analyses

The background characteristics of study participants were described using frequency and percentage distribution. The course of stunting was assessed using cross-sectional prevalence (point prevalence in each round), longitudinal prevalence (the mean probability of being stunted across all rounds of the survey) and annual transition probabilities, mean sojourn-time with 95% confidence interval.[15,20] The confidence interval of the cross-sectional and longitudinal prevalence was adjusted for clustering effect at sentinel level using the 'svy' command in Stata (version 16.0). Further, annual transition probabilities (probability of transitioning from one-stunting state to another in next one year), mean sojourn time (average time spent at a particular state before transitioning to adjacent state) of stunting and transition specific risk factors were estimated through transition intensity matrix using a multistate time-homogeneous continuous-time Markov model (Supplementary Appendix).[21] A three-stunting state (not stunted, moderately stunted, and severely stunted) was constructed for Markov model with the underlying two-step transition (i.e. from the current state to adjacent state) assumption where transitioning to the subsequent states depends only on the current state (Fig. 1). Using Marshall and Jones (1995) method, the hazard ratio for each potential risk factor was estimated for favorable transitions (Moderately stunted to Not Stunted and Severely stunted to Moderately stunted) and unfavorable transitions (not stunted to moderately stunted and moderately stunted to severely stunted).[22] The flexibility of the model allowed for adjusting both time-varying and time-independent factors. For the multistate Markov model, 'msm' R package was used.[21] Further, the favorable and unfavorable transitions across various rounds (proxy for age-group) were assessed by constructing the alluvial plot using '*easyalluvial*' and '*ggplot2*' of R package.[23,24] The 95% confidence interval were

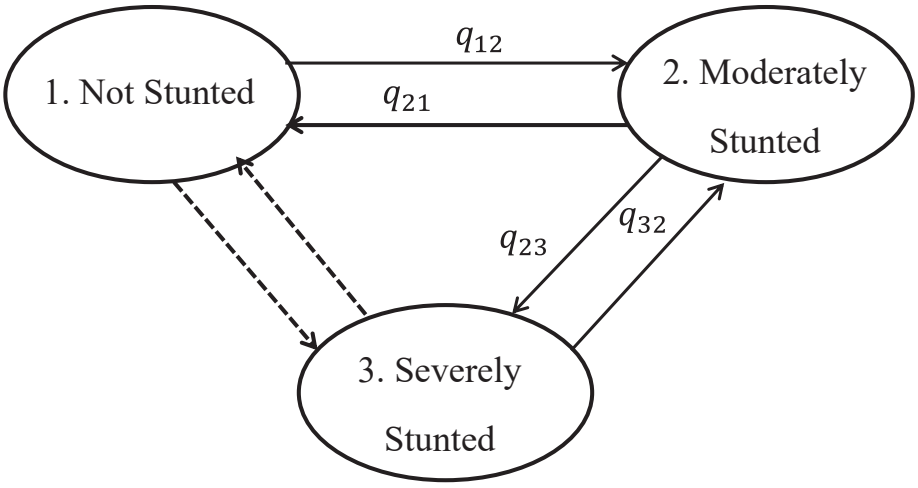


Fig. 1. Transition model of various stunting states ("q" refers to "Transition Intensity").

calculated using the normal approximation. The *P*-value of the hazard ratio was estimated using the method proposed by Altman and Bland.[25–27]

Results

The final sample included 1827 children from the younger cohort with complete follow-up data on the outcome variable and covariates in all 5 survey rounds. The final included sample (n = 1827) was comparable with the excluded sample (n = 184) for most of the sociodemographic characteristics except for the place of residence and religion (Supplementary Table S1). The sociodemographic characteristics of the final sample at the start and end of the longitudinal survey are presented in Table 1. The majority of the children in the sample were boys, and the boy-to-girl ratio (1.18) was similar across rounds 1 and 5. About three-fourth of the cohort were from rural regions, and about one-third belonged to poorer wealth tertile. A significant rural to urban transition and a decrease in household size were observed between rounds 1 and Round 5. The cross-sectional prevalence of severe stunting decreased from 10.45% [191/1827] (95% confidence interval [CI]: 7.8%, 13.7%) to 5.31% [97/1827] (95% CI: 3.8%, 7.3%), while moderate stunting increased from 19.87% [363/1827] (95% CI: 16.3%,23.9%) to 21.73% [397/1827] (95% CI: 18.4%, 24.9%) between Round 1 and Round 5. The mean probability (Longitudinal prevalence) of being not stunted, moderately and severely stunted across all rounds of survey were 70.1% (95% CI: 65.4%,74.3%), 23.45% (95% CI: 20.5%, 26.7%) and 6.50% (95% CI: 5.16%,8.2%) respectively.

The stunting trajectories for the cohort from ages 1 to 15 are shown in Fig. 2 and supplementary Fig. S2-S7. Until age 15, 44.28% [809/1827] (42.65% [421/987] in boys; 46.19% [388/840] in girls) of the children remained not stunted at all time-points (bottom bar without any transitions), 2.03% [37/1827] (2.23% [22/987] in boys; 1.78% [15/840] in girls) remained moderately stunted at all time-points (middle bar without any transitions) and 0.88% [16/1827] (1.11% [11/987] in boys; 0.60% [5/840] in girls) remained severely stunted at all time-points (top bar without any transitions). Among children who were not stunted by age 1, 22.70% [289/1273] (23.26% [154/662] in boys; 22.09% [135/611] in girls) had faltered between ages 1 and 5, 4.88% [48/984] (4.92% [25/508] in boys; 4.83% [23/

Table 1
Sociodemographic characteristics of study cohort at the beginning and end of the longitudinal survey.

Variables	Categories	Round 1 (2002–03)	Round 5 (2016–17)	$\chi 2$ (*P* value)
		Frequency (Percentage)	Frequency (Percentage)	
Sex	Girl	840 (45.98)	840 (45.98)	–
	Boy	987 (54.02)	987 (54.02)	
Place of residence	Urban	446 (24.41)	546 (29.89)	13.84 (<0.001)
	Rural	1381 (75.59)	1281 (70.11)	
Religion	Hindu	1604 (87.79)	1604 (87.79)	–
	Others	223 (12.21)	223 (12.21)	
Social group	Others	1224 (66.99)	1224 (66.99)	–
	SC/ST	603 (33.01)	603 (33.01)	
Mothers' education at birth	No education	1129 (61.79)	1129 (61.79)	–
	Up to grade-5	183 (10.02)	183 (10.02)	
	More than grade-5	515 (28.19)	515 (28.19)	
Household size	Up to 4	771 (42.20)	947 (51.83)	34.03 (<0.001)
	5 or more	1056 (57.80)	880 (48.17)	
Wealth Index	Poor	604 (33.06)	619 (33.88)	0.49 (0.779)
	Middle	606 (33.17)	587 (32.13)	
	Rich	617 (33.77)	621 (33.99)	
Stunting	Not stunted	1273 (69.68)	1333 (72.96)	33.58 (<0.001)
	Moderately stunted	363 (19.87)	397 (21.73)	
	Severely stunted	191 (10.45)	97 (5.31)	
Total		1827 (100)	1827 (100)	

SC/ST denote Scheduled Caste/Scheduled Tribe (a proxy for socio-economically disadvantaged).

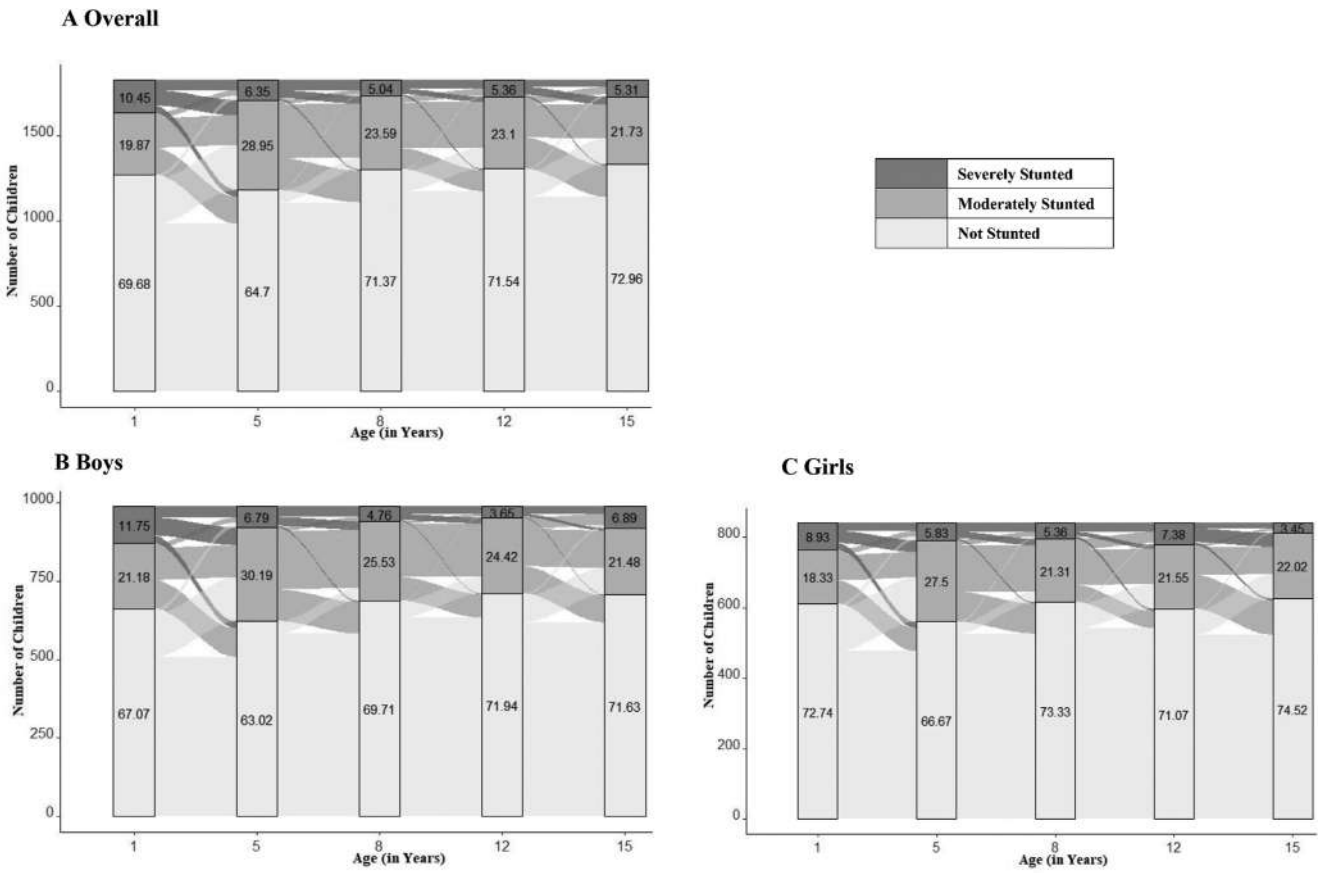


Fig. 2. Alluvial plot showing transitions between stunting states for children from age 1 through age 15 (figures inside the bar represents percentage).

476] in girls) had between ages 5 and 8, 8.07% [85/1053] (7.16% [39/545] in boys; 9.06% [46/508] in girls) had between ages 8 and 12, 10.69% [111/1038] (10.85% [59/544] in boys; 10.53% [52/494] in girls) had between ages 12 and 15. Among the children who were not stunted at either age 1 or age 5, 4.88% [48/984] faltered between the ages of 5 and 8 (4.92% [25/508] in boys; 4.83% [23/476] in girls), 5.99% [59/984] faltered between the ages of 8 and 12 (5.32% [27/508] in boys; 6.72% [32/476] in girls), 6.91% [68/984] faltered between the ages of 12 and 15 (6.89% [35/508] in boys; 6.93% [33/476] in girls).

Among children who were stunted by age 1, 13.13% [26/198] (13.16% [15/114] in boys; 13.10% [11/84] in girls) had faltered between ages 5 and 8, 17.53% [44/251] (11.19% [16/143] in boys; 25.93% [28/108] in girls) faltered between ages 8 and 12, 20.44% [55/269] (19.88% [33/166] in boys; 21.36% [22/103] in girls) faltered between ages 12 and 15. Among children who were stunted by age 1, 35.74% [198/554] (35.08% [114/325] in boys; 36.68% [84/229] in girls) had recovered between ages 1 and 5, 22.19% [79/356] (20.85% [44/211] in boys; 24.14% [35/145] in girls) had recovered between ages 5 and 8, 20.46% [62/303] (21.43% [39/182] in boys; 19.01% [23/121] in girls) had recovered between ages 8 and 12, 28.07% [80/285] (24.53% [39/159] in boys; 32.54% [41/126] in girls) had recovered between ages 12 and 15. Overall, the stunting trajectory had shown gender differential where more faltering to severe stunting and lower recovery to the normal state was observed among girls between 8 and 12 years and among boys between 12 and 15 years.

Overall, the mean sojourn-time derived from the transition intensity matrix for not stunted, moderately stunted, and severely stunted was 14.78 years (95% CI: 6.56, 35.83), 3.31 years (95% CI: 1.42, 6.48), and 6.47 years (95% CI: 1.40, 29.69), respectively, indicating a relatively concise state for moderate stunting. Further analysis of annual transition probabilities showed a higher tendency for the moderately stunted state to shift toward a not stunted state than the severely stunted state (Table 2). A similar trend was observed across all covariates except in the 1–5 year age group. Compared with their respective counterparts, being a girl child, belonging to an urban area, non-SC/ST social group, richer wealth tertile, and whose mother's education at birth more than primary-level had a higher annual probability for a favorable transition from moderate to not stunted state. Notably, a girl child and a child whose mother's education at birth was more than primary-level also had a higher annual probability for an unfavorable transition from moderate to severely stunted state. For a non-stunted child, being a girl, belonging to the urban area and SC/ST social group had a higher annual probability for an unfavorable transition to the moderately stunted state.

The risk factor analysis using the multistate Markov model showed the child's gender, place of residence, mother's education at birth, and household size as significant determinants for stunting transitions (Table 3). Children of educated mothers (>primary schooling) and children from a large household had a lesser likelihood of unfavorable transition from not stunted to moderately

Table 2
Annual transition probabilities for various sociodemographic factors.

Covariates		Current State	Transitioning state		
			Not Stunted	Moderately stunted	Severely stunted
			Probability (95% CI)	Probability (95% CI)	Probability (95% CI)
Age	Overall (1–15 years)	Not stunted	0.94 (0.87–0.97)	0.06 (0.02–0.12)	0.00 (0.00–0.01)
		Moderately stunted	0.20 (0.09–0.40)	0.75 (0.52–0.86)	0.05 (0.01–0.23)
		Severely stunted	0.02 (0.00–0.08)	0.12 (0.03–0.47)	0.86 (0.46–0.97)
	1–5 years	Not stunted	0.80 (0.44–0.95)	0.13 (0.01–0.37)	0.08 (0.00–0.42)
		Moderately stunted	0.19 (0.01–0.58)	0.30 (0.01–0.81)	0.51 (0.02–0.97)
		Severely stunted	0.01 (0.00–0.10)	0.06 (0.00–0.26)	0.93 (0.65–1.00)
	5–8 years	Not stunted	0.99 (0.89–1.00)	0.01 (0.00–0.11)	0.00 (0.00–0.01)
		Moderately stunted	0.41 (0.11–0.88)	0.57 (0.09–0.85)	0.02 (0.00–0.44)
		Severely stunted	0.21 (0.01–0.76)	0.54 (0.02–0.83)	0.25 (0.00–0.97)
	8–12 years	Not stunted	0.99 (0.93–1.00)	0.02 (0.00–0.07)	0.00 (0.00–0.00)
		Moderately stunted	0.07 (0.01–0.37)	0.93 (0.54–0.98)	0.01 (0.00–0.22)
		Severely stunted	0.00 (0.00–0.05)	0.03 (0.00–0.68)	0.97 (0.25–1.00)
	12–15 years	Not stunted	0.96 (0.83–0.99)	0.04 (0.01–0.17)	0.00 (0.00–0.01)
		Moderately stunted	0.11 (0.02–0.43)	0.88 (0.52–0.97)	0.01 (0.00–0.15)
		Severely stunted	0.01 (0.00–0.15)	0.09 (0.00–0.89)	0.90 (0.03–1.00)
Gender	Boy	Not stunted	0.94 (0.87–0.97)	0.06 (0.03–0.12)	0.00 (0.00–0.02)
		Moderately stunted	0.23 (0.11–0.43)	0.70 (0.46–0.83)	0.07 (0.01–0.29)
		Severely stunted	0.02 (0.01–0.09)	0.15 (0.04–0.46)	0.83 (0.47–0.96)
	Girl	Not stunted	0.93 (0.86–0.97)	0.06 (0.03–0.13)	0.00 (0.00–0.02)
		Moderately stunted	0.27 (0.14–0.50)	0.64 (0.37–0.79)	0.09 (0.02–0.33)
		Severely stunted	0.03 (0.01–0.12)	0.18 (0.05–0.48)	0.79 (0.41–0.94)
Place of residence	Urban	Not stunted	0.93 (0.87–0.96)	0.07 (0.04–0.13)	0.00 (0.00–0.01)
		Moderately stunted	0.14 (0.07–0.26)	0.82 (0.66–0.89)	0.04 (0.01–0.15)
		Severely-stunted	0.01 (0.00–0.04)	0.14 (0.05–0.35)	0.85 (0.62–0.95)
	Rural	Not stunted	0.92 (0.87–0.96)	0.08 (0.04–0.13)	0.00 (0.00–0.01)
		Moderately stunted	0.10 (0.06–0.17)	0.87 (0.78–0.92)	0.03 (0.01–0.10)
		Severely stunted	0.01 (0.00–0.02)	0.15 (0.06–0.36)	0.85 (0.62–0.94)
Religion	Hindu	Not stunted	0.94 (0.89–0.97)	0.06 (0.03–0.11)	0.00 (0.00–0.01)
		Moderately stunted	0.19 (0.10–0.34)	0.77 (0.59–0.87)	0.04 (0.01–0.17)
		Severely stunted	0.02 (0.00–0.07)	0.16 (0.04–0.48)	0.82 (0.45–0.95)
	Others	Not stunted	0.94 (0.89–0.97)	0.06 (0.03–0.11)	0.00 (0.00–0.01)
		Moderately stunted	0.17 (0.09–0.31)	0.79 (0.63–0.88)	0.04 (0.01–0.15)
		Severely stunted	0.02 (0.01–0.09)	0.21 (0.06–0.58)	0.77 (0.35–0.94)
Social group	SC/ST	Not stunted	0.93 (0.86–0.97)	0.07 (0.03–0.13)	0.00 (0.00–0.01)
		Moderately stunted	0.19 (0.09–0.36)	0.74 (0.52–0.86)	0.06 (0.01–0.25)
		Severely stunted	0.01 (0.00–0.06)	0.10 (0.03–0.37)	0.88 (0.59–0.97)
	Others	Not stunted	0.94 (0.87–0.97)	0.06 (0.03–0.12)	0.00 (0.00–0.01)
		Moderately stunted	0.20 (0.09–0.36)	0.75 (0.54–0.86)	0.06 (0.01–0.23)
		Severely stunted	0.01 (0.00–0.06)	0.11 (0.03–0.40)	0.87 (0.53–0.97)
Mother's education at birth	No formal education	Not stunted	0.95 (0.90–0.98)	0.05 (0.02–0.10)	0.00 (0.00–0.01)
		Moderately stunted	0.21 (0.10–0.40)	0.73 (0.52–0.85)	0.06 (0.01–0.23)
		Severely stunted	0.02 (0.00–0.09)	0.14 (0.03–0.47)	0.84 (0.46–0.97)
	Up to 5	Not stunted	0.95 (0.91–0.98)	0.04 (0.02–0.09)	0.00 (0.00–0.01)
		Moderately stunted	0.22 (0.11–0.42)	0.72 (0.48–0.84)	0.06 (0.01–0.25)
		Severely stunted	0.02 (0.01–0.09)	0.15 (0.04–0.44)	0.83 (0.48–0.96)
	More than 5	Not stunted	0.96 (0.91–0.98)	0.04 (0.02–0.08)	0.00 (0.00–0.01)
		Moderately stunted	0.24 (0.11–0.43)	0.70 (0.45–0.84)	0.07 (0.01–0.26)
		Severely stunted	0.03 (0.01–0.10)	0.17 (0.04–0.47)	0.81 (0.44–0.95)
Household size	Up to 4	Not stunted	0.95 (0.90–0.98)	0.05 (0.02–0.10)	0.00 (0.00–0.01)
		Moderately stunted	0.20 (0.10–0.38)	0.75 (0.53–0.87)	0.05 (0.01–0.26)
		Severely stunted	0.01 (0.00–0.06)	0.09 (0.02–0.39)	0.90 (0.56–0.98)
	More than 4	Not stunted	0.96 (0.91–0.98)	0.04 (0.02–0.09)	0.00 (0.00–0.01)
		Moderately stunted	0.21 (0.10–0.40)	0.75 (0.52–0.87)	0.05 (0.01–0.21)
		Severely stunted	0.01 (0.00–0.04)	0.07 (0.02–0.27)	0.92 (0.70–0.98)
Wealth Index	Poor	Not stunted	0.95 (0.89–0.98)	0.05 (0.02–0.11)	0.00 (0.00–0.01)
		Moderately stunted	0.21 (0.10–0.39)	0.75 (0.53–0.86)	0.05 (0.01–0.19)
		Severely stunted	0.02 (0.00–0.08)	0.14 (0.04–0.46)	0.84 (0.46–0.96)
	Middle	Not stunted	0.95 (0.91–0.98)	0.05 (0.02–0.09)	0.00 (0.00–0.01)
		Moderately stunted	0.21 (0.10–0.40)	0.75 (0.55–0.86)	0.04 (0.01–0.16)
		Severely stunted	0.02 (0.01–0.09)	0.17 (0.04–0.48)	0.81 (0.45–0.95)
	Rich	Not stunted	0.96 (0.91–0.98)	0.04 (0.02–0.08)	0.00 (0.00–0.01)
		Moderately stunted	0.22 (0.11–0.40)	0.75 (0.56–0.86)	0.04 (0.01–0.13)
		Severely stunted	0.03 (0.01–0.10)	0.19 (0.06–0.53)	0.78 (0.38–0.93)

SC/ST denote Scheduled Caste/Scheduled Tribe (a proxy for socio-economically disadvantaged). CI, confidence interval.

stunted. Compared with boys, girl child had 40% excess likelihood (Hazard Ratio: 1.40; 95% CI 1.00 to 1.95) for an unfavorable transition from moderately stunted to severely stunted and 19% excess likelihood (Hazard Ratio: 1.19; 95% CI 1.01 to 1.40) for a favorable transition from moderately stunted to not-stunted state. Children from the rural area and children of mothers with no formal education had a lesser likelihood for a favorable transition from moderately stunted to not-stunted state.

Table 3
Hazard ratios from multistate Markov model fitted separately for covariates of transitions between not stunted, moderately stunted, and severely stunted across children in India.

Covariates		Not stunted to Moderately stunted		Moderately stunted to Severely stunted		Moderately stunted to Not Stunted		Severely stunted to Moderately stunted	
		Hazard Ratio (95% CI)	*P* value	Hazard Ratio (95% CI)	*P* value	Hazard Ratio (95% CI)	*P* value	Hazard Ratio (95% CI)	*P* value
Gender	Boy	Ref		Ref		Ref		Ref	
	Girl	1.08 (0.92–1.28)	0.367	1.40 (1.01–1.95)	**0.044**	1.19 (1.01–1.40)	**0.037**	1.28 (0.97–1.68)	0.084
Place of residence	Urban	Ref		Ref		Ref		Ref	
	Rural	1.08 (0.83–1.41)	0.566	0.76 (0.44–1.32)	0.338	0.68 (0.54–0.86)	**0.001**	0.99 (0.60–1.63)	0.976
Religion	Hindu	Ref		Ref		Ref		Ref	
	Others	0.96 (0.73–1.26)	0.788	0.85 (0.47–1.53)	0.598	0.90 (0.69–1.16)	0.417	1.32 (0.81–2.15)	0.262
Social Group	Others	Ref		Ref		Ref		Ref	
	SC/ST	1.11 (0.93–1.34)	0.248	1.11 (0.79–1.56)	0.559	1.00 (0.84–1.18)	0.958	0.93 (0.70–1.22)	0.593
Mother's education at birth	No formal education	Ref		Ref		Ref		Ref	
	Up to grade 5	1.21 (0.91–1.61)	0.186	1.28 (0.73–2.24)	0.397	1.37 (1.04–1.80)	**0.023**	1.08 (0.68–1.70)	0.769
	> grade 5	0.78 (0.62–0.98)	**0.036**	1.24 (0.75–2.04)	0.404	1.10 (0.88–1.36)	0.426	1.34 (0.89–2.01)	0.165
Household size	Up to 4	Ref		Ref		Ref		Ref	
	5 or more	0.84 (0.71–0.99)	**0.036**	0.94 (0.67–1.32)	0.748	1.01 (0.86–1.19)	0.905	0.72 (0.55–0.95)	**0.019**
Wealth Index	Poor	Ref		Ref		Ref		Ref	
	Middle	1.01 (0.82–1.23)	0.966	0.89 (0.60–1.30)	0.546	1.05 (0.87–1.28)	0.610	1.28 (0.93–1.75)	0.126
	Rich	0.78 (0.59–1.02)	0.126	0.82 (0.47–1.42)	0.233	1.06 (0.82–1.35)	0.760	1.30 (0.80–2.11)	0.116

SC/ST denote Scheduled Caste/Scheduled Tribe (a proxy for socio-economically disadvantaged). CI, confidence interval.

Discussion

The present study provides several contributions to child and adolescent malnutrition in India by investigating the course of childhood stunting using YLS longitudinal data. To the best of our knowledge, this is the first study to provide comprehensive analysis and transition probabilities of stunting for Indian children using a time-homogeneous continuous-time multistate Markov transition model. In emphasizing the importance of adolescent nutrition, this longitudinal analysis had found considerable transitions across stunting categories throughout childhood and had provided useful insights to promote child growth beyond 1000 days. The study highlighted the vulnerability of moderately stunted under-5 children and girl children (especially aged 8–12 years) for more faltering than recovery. These findings have policy and programmatic implications for strengthening nutrition interventions for children and adolescents in India.

The longitudinal prevalence of 29.95% reported for stunting in the current study falls under the higher threshold of public health significance.[28] Within the complex growth trajectories of both growth recovery and faltering, the study did not find a substantial girl disadvantage for recovery at various time points till 8 years of age. With gender inequities continuing to exist in India, further examination is required to explore the possible role of an inherent biological difference to cope or possible survivor bias (where the surviving girl child are better cared) to support this finding.[14,29] At the same time, the study also documented a more faltering than recovery for girl children between 8 and 12 years. This crucial finding emphasizes 8–12 years as another critical period of growth, especially for girl children in India. Although an increased need from the early entry of puberty for girl child during this period can be attributed, further research is required to understand the contextual factors like gender discrimination, gender disparity, and intrahousehold discrimination that made those needs unmet for an adolescent girl child in India.[30–32]

In contrast to severe stunting, moderate stunting is associated with a larger proportion of nutrition-related deaths.[13] Mean sojourn-time and annual transition probabilities estimated from the current study had shown moderately stunted state as a concise state with more tendency for recovery than faltering in all age-groups except in 1–5 years. This would imply a need for differential policy and programmatic approaches to address moderate and severe stunting in India, especially in under-5 children.[33] Longitudinal prevalence of 23.45% for moderate stunting across the age groups further emphasizes a need for integrating periodic monitoring of nutritional status (to differentiate severe and moderate stunting) into the ongoing nutrition programs and strengthening its scope beyond under-5 children to reduce the burden of malnutrition in India.[33]

Further, the stunting trajectory had shown gender differential where more faltering to severe stunting and lower recovery to the normal state was observed among girls between 8 and 12 years and among boys between 12 and 15 years. Although this finding could be explained by differential entry into puberty, exploring life stages beyond the first 1000 days, especially preadolescence as another window of opportunity to prevent malnutrition in adolescence, has been emphasized.[34] The risk analysis had shown moderately stunted girl child to have a higher risk for faltering than recovery as compared with the boy child. Without seeking to undermine the interventions for severe stunting, this highlights moderately stunted girl children as a highly vulnerable group emphasizing the need for a gender-responsive approach to promoting favorable growth transition among girl children.

The factors that contribute to stunting and its course are multiple and interlinked at household, socioeconomic and

environmental level.[7] The study found a child's gender, place of residence, mother's education at birth, and household size as significant determinants for stunting transitions. With the simultaneous influence of different risk factors at various levels and their linkages, an integrated multilevel approach has been emphasized to address the determinants of undernutrition.[35,36] However, it is important to note that the coverage of adolescent nutrition interventions have remained low and uneven in India.[37] Thus, strengthening and expanding the integrated multilevel approach to cover preadolescence and adolescence would be crucial to achieving the goals set by the National Nutrition Mission.[38]

The study has several strengths. First, the study had used a time-homogeneous continuous-time Markov model over the standard multistate model to adjust for irregularities in the follow-up data and lack of information on the exact time of transition in the dataset. Second, the study used data from a relatively large cohort of children in India who were followed longitudinally from birth to age 15. There are very few child cohorts in India who were followed from birth into adolescence. Third, the loss to followup is relatively low for the younger cohort. This enables the study findings to be more valid internally for the given context. Fourth, the study used height (stunting) as an outcome measure, which is measured objectively in contrast to other measures like food intake, dietary diversity, etc., that are often reported by caregivers. Also, height is often regarded as a more acceptable measure to investigate gender differences in child nutrition, especially in the absence of objective measures of food consumption and other intrahousehold allocation of resources.[39] Further, stunting was preferred over the other outcome variables of malnutrition (viz. wasting, BMI for age) due to their limitations of the operational definition and utility across various age groups.[18]

The study has the following limitations. The study used data from YLS, which surveyed children in the state of Andhra Pradesh (later split into Andhra Pradesh and Telangana) and hence may not be nationally representative. Despite this, the sample covered a diversity of children in terms of wealth, consumption, health, nutrition, and education similar to national datasets and thus was more robust for understanding longitudinal dynamics of stunting across the life course of children.[16] The study did not use change in z-scores of attained heights (ΔHAZ), which would have better reflected a dynamic assessment of growth.[40] However, it is important to note that the dynamic range of growth related to pathology might not be actually linear. Further, it was widely acknowledged that the categorization of HAZ scores used in the current study had more relevance for public health surveillance and policy decisions. We could not explore several risk factors like infections, immunizations, birth order, micronutrient deficiencies, siblings' gender composition as these were neither complete nor were captured during the survey. Although it was found that birth order and siblings' gender composition do not influence children's height, the effects of other determinants remain unknown.[41]

Conclusion

This comprehensive analysis from India had found preadolescence as an additional critical window for nutritional intervention to influence policy for promoting better nutrition within the complex growth trajectory. Further, the gender differentials for faltering and recovery re-emphasize the importance of strengthening gender-responsive approaches in the National Nutrition Strategy of India to tackle malnutrition. Children in India receive several nutrition services through various programs. Differential approach for identification and management of moderate and severe stunting although periodic monitoring of nutritional status in such programs thus becomes crucial to prevent undernutrition across the life cycle. With a fifth of adolescents living in India, a coordinated, multisectoral, age-appropriate and gender-responsive approach is required to take India closer to meeting SDG-2.

Author statements

Acknowledgments

The data used in this publication come from Young Lives, a 15-year study of the changing nature of childhood poverty in Ethiopia, India (Andhra Pradesh and Telangana), Peru, and Vietnam (www.younglives.org.uk). Young Lives is core-funded by UK aid from the Department for International Development (now the Foreign, Commonwealth & Development Office), United Kingdom. The views expressed here are those of the author(s). They are not necessarily those of Young Lives, the University of Oxford, DFID, or other funders.

Ethical approval

Not required. The data used in this publication come from the YLS study. Details of ethics approval for the YLS study can be obtained from www.younglives.org.uk/content/research-ethics. The YLS data set is freely available and can be accessed for further use after the registration process. The results obtained in the current study are based on the secondary analysis of the existing YLS dataset and do not contain the respondent's name or any other identifiers.

Funding

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Competing interests

None declared.

Authors contribution

Sumit Kumar Das: Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Writing- Original draft preparation, review, and editing. **Ajit Deo Burma**: Conceptualization, Validation, Writing- Original draft preparation, review, and editing. **Vinayak Mishra**: Conceptualization, Validation, Writing- Original draft preparation, review, and editing. **Senthil Amudhan**: Conceptualization, Methodology, Investigation, Supervision, Validation, Writing – Original draft preparation, review, and editing. **Payel Mahapatra**: Resources, Writing- Original draft review and editing. **Ashi Ashok**: Resources, Writing- Original draft review and editing. **Mariamma Philip**: Investigation, Supervision, Validation, Writing – review and editing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2021.10.010.

References

1. India State-Level Disease Burden Initiative Malnutrition Collaborators. The burden of child and maternal malnutrition and trends in its indicators in the states of India: the Global Burden of Disease Study 1990-2017. *Lancet Child Adolesc Health* 2019 Dec;**3**(12):855–70.
2. Development Initiatives. *Global Nutrition Report 2018: shining a light to spur action on nutrition*. Bristol: UK: Development Initiatives; 2018.
3. Ministry of Health and Family Welfare, Govt. of India. *UNICEF, population council. Comprehensive national nutrition survey (CNNS) national report*. 2019 [Internet]. New Delhi [cited 2020 Dec 16]. Available from: https://nhm.gov.in/index1.php?lang=1&level=2&sublinkid=1332&lid=713.
4. Horton S, Steckel RH. Malnutrition: global economic losses attributable to malnutrition 1900–2000 and projections to 2050. In: Bjørn Lomborg, editor. *How much have global problems cost the world?* Cambridge: Cambridge University Press; 2013. p. 247–72.
5. Asiki G, Newton R, Marions L, Kamali A, Smedman L. The effect of childhood stunting and wasting on adolescent cardiovascular diseases risk and educational achievement in rural Uganda: a retrospective cohort study. *Glob Health Action* 2019 Jun 24;**12**(1) [Internet] [cited 2020 Dec 26] Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6598535/.
6. Richter LM, Daelmans B, Lombardi J, Heymann J, Boo FL, Behrman JR, et al. Investing in the foundation of sustainable development: pathways to scale up for early childhood development. *Lancet* 2017 Jan 7;**389**(10064):103–18.
7. Black RE, Victora CG, Walker SP, Bhutta ZA, Christian P, de Onis M, et al. Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet* 2013 Aug 3;**382**(9890):427–51.
8. Shrimpton R, Victora CG, de Onis M, Lima RC, Blössner M, Clugston G. Worldwide timing of growth faltering: implications for nutritional interventions. *Pediatrics* 2001 May;**107**(5):E75.
9. Prentice AM, Ward KA, Goldberg GR, Jarjou LM, Moore SE, Fulford AJ, et al. Critical windows for nutritional interventions against stunting. *Am J Clin Nutr* 2013 May;**97**(5):911–8.
10. Gausman J, Kim R, Subramanian SV. Stunting trajectories from post-infancy to adolescence in Ethiopia, India, Peru, and Vietnam. *Matern Child Nutr* 2019 Oct;**15**(4):e12835.
11. Lundeen EA, Behrman JR, Crookston BT, Dearden KA, Engle P, Georgiadis A, et al. Growth faltering and recovery in children aged 1-8 years in four low- and middle-income countries: Young Lives. *Publ Health Nutr* 2014 Sep;**17**(9):2131–7.
12. Fink G, Rockers PC. Childhood growth, schooling, and cognitive development: further evidence from the Young Lives study. *Am J Clin Nutr* 2014 Jul;**100**(1):182–8.
13. Shoham J, Duffield A. Proceedings of the world health organization/UNICEF/world food programme/united nations high commissioner for refugees consultation on the management of moderate malnutrition in children under 5 years of age. *Food Nutr Bull* 2009 Sep;**30**(3 Suppl):S464–74.
14. Krishna A, Oh J, Perkins JM, Lee H-Y, Heo J, Lee J-K, et al. Is there a female disadvantage in child undernutrition in South India?: exploring gender differences in height in infancy, childhood, and adolescence in Andhra Pradesh and Telangana. *Am J Hum Biol* 2019 Jan;**31**(1):e23153.
15. Astatkie A. Dynamics of stunting from childhood to youthhood in Ethiopia: evidence from the Young Lives panel data. *PLoS One* 2020;**15**(2):e0229011.
16. Young Lives-India. *Young Lives survey design and sampling (Round 5): India*. Oxford: Oxford Department of International Development; 2017 [Internet] [cited 2020 Dec 18]. Available from: https://younglives-india.org/sites/www.younglives-india.org/files/INDIA-SurveyDesign-Factsheet-Oct17_0.pdf.
17. World Health Organization. *WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age : methods and development*. 2006. Geneva.
18. de Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ* 2007 Sep;**85**(9):660–7.
19. Briones KJ. How many rooms are there in your house?. In: *Constructing the Young Lives wealth index*. Oxford: Young Lives,Oxford Department of International Development (ODID); 2017 Nov. p. 1–18. Report No.: Technical Note 43.
20. Schmidt W-P, Genser B, Luby SP, Chalabi Z. Estimating the effect of recurrent infectious diseases on nutritional status: sampling frequency, sample-size, and bias. *J Health Popul Nutr* 2011 Aug;**29**(4):317–26.
21. Jackson C. Multi-state models for panel data: the msm package for R. *J Stat Software* 2011 Jan 4;**38**(1):1–28.
22. Marshall G, Jones RH. Multi-state models and diabetic retinopathy. *Stat Med* 1995 Sep 30;**14**(18):1975–83.
23. Koneswarakantha B. *easyalluvial: generate alluvial plots with a single line of code (Version 0.2.3) [R-Packages]. https://CRAN.R-project.org/package=easyalluvial*. 2020 [Internet] [cited 2020 Dec 18]. Available from: https://CRAN.R-project.org/package=easyalluvial.
24. Wickham H, Chang W, Henry L, Pedersen TL, Takahashi K, Wilke C, et al. *ggplot2: create elegant data visualisations using the grammar of graphics (Version 3.3.2) [R-Packages]. https://CRAN.R-project.org/package=ggplot2*. 2020 [Internet] [cited 2020 Dec 18]. Available from: https://CRAN.R-project.org/package=ggplot2.
25. Altman DG, Bland JM. How to obtain the P value from a confidence interval. *BMJ* 2011;**343**:d2304.
26. Meitei WB, Ladusingh L. Transition specific risk factors affecting the lifestyle disease progression from diabetes to hypertension in India. *Health* 2019 Aug 13;**11**(8):1055–71.
27. Moreira C, Meira-Machado L, Fonseca MJ, Santos AC. A multistate model for analyzing transitions between body mass index categories during childhood: the generation XXI birth cohort study. *Am J Epidemiol* 2019 Feb 1;**188**(2):305–13.
28. de Onis M, Borghi E, Arimond M, Webb P, Croft T, Saha K, et al. Prevalence thresholds for wasting, overweight and stunting in children under 5 years. *Publ Health Nutr* 2019 Jan;**22**(1):175–9.
29. Marcoux A. Sex differentials in undernutrition: a look at survey evidence. *Popul Dev Rev* 2002;**28**(2):275–84.
30. Soliman A, De Sanctis V, Elalaily R. Nutrition and pubertal development. *Indian J Endocrinol Metab* 2014 Nov;**18**(Suppl 1):S39–47.
31. Rajan S, Morgan SP. Selective versus generalized gender bias in childhood health and nutrition: evidence from India. *Popul Dev Rev* 2018 Jun;**44**(2):231–55.
32. Jose S. Decomposition of gender differential in malnutrition in Indian children. *J Soc Econ Dev* 2017 Oct 1;**19**(2):299–322.
33. Banerjee K, Dwivedi LK. Disparity in childhood stunting in India: relative importance of community-level nutrition and sanitary practices. *PLoS One* 2020;**15**(9):e0238364.
34. Lassi Z, Moin A, Bhutta Z. Nutrition in middle childhood and adolescence. In: Bundy DAP, de Silva N, Horton S, Jamison DT, Patton GC, editors. *Child and adolescent health and development*. 3rd ed. Washington (DC): The International Bank for Reconstruction and Development / The World Bank; 2017 [Internet] [cited 2021 Jun 17]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK525242/.
35. Khan J, Das SK. The burden of anthropometric failure and child mortality in India. *Sci Rep* 2020 Dec 2;**10**(1):20991.
36. Subramanian SV, Mejía-Guevara I, Krishna A. Rethinking policy perspectives on childhood stunting: time to formulate a structural and multifactorial strategy. *Matern Child Nutr* 2016 May 17;**12**(Suppl 1):219–36.
37. Gupta AK, Santhya KG. Proximal and contextual correlates of childhood stunting in India: a geo-spatial analysis. *PLoS One* 2020 Aug 20;**15**(8):e0237661.
38. *POSHAN abhiyaan | NITI aayog* [Internet]. [cited 2021 Jun 17]. Available from: http://niti.gov.in/poshan-abhiyaan.
39. Krishna A, Fink G, Berkman LF, Subramanian SV. Short- and long-run associations between birth weight and children's height. *Econ Hum Biol* 2016 May;**21**:156–66.
40. Wit JM, Himes JH, van Buuren S, Denno DM, Suchdev PS. Practical application of linear growth measurements in clinical research in low- and middle-income countries. *Horm Res Paediatr* 2017;**88**(1):79–90.
41. Krishna A, Mejía-Guevara I, McGovern M, Aguayo VM, Subramanian SV. Trends in inequalities in child stunting in South Asia. *Matern Child Nutr* 2018 Nov;**14**(Suppl 4):e12517.

Public Health 202 (2022) 113–120

Contents lists available at ScienceDirect

Public Health

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

Original Research

Early-life exposure to famine and the risk of general and abdominal obesity in adulthood: a 22-year cohort study

D. Liu [a, c], J. Yang [b, c], S. Wang [a, *]

[a] Department of Endocrinology, Tianjin Fifth Central Hospital, Tianjin, 300450, China
[b] Department of Orthopedics, Tianjin Fifth Central Hospital, Tianjin, 300450, China

ARTICLE INFO

Article history:
Received 24 June 2021
Received in revised form
8 November 2021
Accepted 19 November 2021
Available online 20 December 2021

Keywords:
Famine
Overweight
Obesity
Abdominal obesity
China

ABSTRACT

Objectives: The aim of this study was to investigate how early-life exposure to famine affected the development of overweight, general obesity and abdominal obesity in Chinese adults.
Study design: This study was a 22-year cohort study.
Methods: Data were derived from the China Health and Nutrition Survey, which is a national prospective cohort study. All participants born between 1949 and 1966 were potentially eligible. Height, weight and waist circumference (WC) were measured by trained healthcare workers. Height and weight were used to calculate body mass index, which was used to define general obesity and WC was used to define abdominal obesity. Exposure to famine was defined using the birth date as follows: no exposure (participants born between 1962 and 1966); fetal exposure (participants born between 1959 and 1961); early childhood exposure (participants born between 1956 and 1958); mid-childhood exposure (participants born between 1953 and 1955); and late childhood exposure (participants born between 1949 and 1952).
Results: In total, 6957 participants were included in this study. Results indicate that exposure to famine was linked to a lower risk of being overweight. Exposure to famine in mid-childhood decreased the risk of general obesity in both males (hazard ratio [HR] 0.485, 95% confidence interval [CI] 0.292–0.807 [$P = 0.005$]) and females (HR 0.426, 95% CI 0.256–0.709 [$P = 0.001$]). Exposure to famine during any period of childhood decreased the risk of abdominal obesity ($P < 0.001$).
Conclusions: Exposure to famine in early childhood decreased the risk of overweight and abdominal obesity in adulthood; however, exposure to famine only had a weak role in the development of general obesity.

Introduction

Obesity is an important public health problem worldwide because of the adverse effects it has on chronic diseases.[1–7] As the economy grows and living standards improve rapidly, overweight and general obesity are becoming an even greater problem. In 2010, overweight and general obesity accounted for approximately 3.8% of disability-adjusted life years, 3.4 million deaths and 3.9% of years of life lost globally.[8] Moreover, 40.5% and 57.0% of US adults suffered from general and abdominal obesity in 2016.[9] In recent decades, the prevalence of obesity increased from 12.6% in 1980 to 30.5% in 2015.[10] The prevalence of abdominal obesity among adults in worldwide reached 47.34% in 2011;[11] therefore, obesity is a major public health issue in China.

Major natural and social events in early life can have adverse effects on health outcomes in adulthood, especially on chronic diseases.[12] The Critical Period Hypothesis suggests that the environment in fetal life has a lifelong impact on body constitution.[13] The Fetal Origins Hypothesis suggests that undernutrition in fetal life has an adverse effect on health outcomes in adulthood.[14–16] Some studies have reported the associations of famine with the risks of overweight,[17] general obesity[16,18] and abdominal obesity.[19] However, the conclusions are controversial and most studies failed to consider the influence of early childhood age and location.[20] As the Chinese Famine was a period between 1959 and 1961, individuals who had been exposed to famine were older than unexposed participants; thus, making it difficult to distinguish the effect of famine exposure from the effect of aging. In addition, the severity

* Corresponding author. Department of Endocrinology, Tianjin Fifth Central Hospital, No 41 Zhejiang Road, Tianjin 300450, China. Tel./fax: +86 022 65665880.
E-mail address: wsrealm@163.com (S. Wang).
[c] These authors contributed equally to this work.

https://doi.org/10.1016/j.puhe.2021.11.014
0033-3506/

of famine varied in different areas. Therefore, age and famine severity are the key confounding factors and should be adjusted for in the study results.

Given the epidemic of obesity in China and the long-term impact of the Chinese Famine, properly determining the association of exposure to famine with the risk of obesity would identify famine survivors who may benefit from targeted obesity management and also provide evidence for essential nutrition intake in early life.

This study was designed to examine the associations of exposure to famine in early life with the development of overweight, general and abdominal obesity in adulthood among Chinese adults.

Methods

Study design and population

Data analyzed in this study were derived from the China Health and Nutrition Survey (CHNS) conducted from 1989 to 2011, which aimed to monitor the health and nutritional status of the population in China. The first wave of the CHNS was conducted in 1989 and subsequently in 1991, 1993, 1997, 2000, 2004, 2006, 2009 and 2011. A multistage, random cluster process was used to select a study sample from nine provinces, which were randomly selected from 31 provinces in China. Counties or neighborhoods in the targeted provinces were stratified by income (low, middle and high) to randomly select four counties or neighborhoods in each province. Then, villages and townships within the targeted counties and neighborhoods were randomly selected. The detailed selection process has been described elsewhere.[21] Since the targeted provinces were representative in geography, economic development, public resources, and health indicators, the study sample was representative of the Chinese population. All individuals born between 1949 and 1966 were potentially eligible for inclusion in the study unless they had missing or abnormal data.

Overweight, general and abdominal obesity

Height, weight and waist circumference (WC) were measured by trained healthcare workers or members of the CHNS following standardized protocols. Participants were requested to remove their shoes and wear lightweight clothing to complete measurements of height and weight using a portable stadiometer and a calibrated beam scale, respectively. A non-elastic tape was used to measure WC in a horizontal plane. The detailed measurements of height, weight, and WC have been described in a previous publication.[11]

Individuals with a body mass index (BMI) $\geq$23 kg/m^2 to <27.5 kg/m^2 were identified as overweight, and those with a BMI $\geq$27.5 kg/m^2 were identified as having general obesity.[22] Males with a WC $\geq$ 90 cm or females with a WC $\geq$ 80 cm were identified as having abdominal obesity.[23]

Definition of famine exposure

The Chinese Famine occurred from 1959 to 1961; thus, the period of famine exposure was defined as being between 1959 and 1961.[24] According to a previous study, famine exposure was defined using birth dates as follows: individuals born between January 1, 1962 and June 31, 1966 were considered to have 'no exposure'; those born between January 1, 1959 and December 31, 1961 were considered to have 'fetal exposure'; those born between January 1, 1956 and December 31, 1958 were considered to have 'early childhood exposure'; those born between January 1, 1953 and December 31, 1955 were considered to have 'mid-childhood exposure'; and those born between January 1, 1949 and December 31, 1952 were considered to have 'late childhood exposure'.[25]

The Chinese Famine led to increased mortality rates, which varied across different provinces; thus, the excess death rate (EDR) was used to assess famine severity in different regions in this study. EDR was calculated as the percentage change of the difference between the highest mortality rate during the Chinese Famine from 1959 to 1961 and the average mortality rate from 1956 to 1958, divided by the average mortality rate from 1956 to 1958. If an EDR of any area was $\geq$100%, it was considered as an area that experienced severe famine; otherwise, it was considered as an area that experienced less severe famine.

Covariates

Data on sex (male or female), ethnicity (Han or others), famine severity (severe or less severe) and living areas (rural or urban) were collected using a valid questionnaire. Smoking status was identified using a question: "Do you smoke cigarettes now (including hand-rolled or device-rolled)?" An answer of 'no' was considered as no smoking and 'yes' was considered as smoking. Similarly, alcohol consumption was identified using a question: "Did you drink beer or any other alcoholic beverage?" An answer of 'no' was considered as no alcohol consumption and 'yes' was considered as alcohol consumption. All participants were asked questions regarding physical activities, which included six activity types: (1) martial arts; (2) gymnastics, dancing and acrobatics; (3) track and field and swimming; (4) soccer, basketball and tennis; (5) badminton and volleyball; and (6) other (e.g. ping pong, Tai Chi, etc.). Individuals participating in at least one of these activities were considered to be physically active. Gross family income was divided into low or high, using a median of 20,000 RMB.

Statistical analyses

Means ± standard deviations (SDs) and frequencies (percentages) were used to describe the distributions of continuous and categorical variables, respectively. Comparisons across different famine exposure groups in all baseline characteristics were conducted using analysis of variance and chi-square tests for continuous and categorical variables, respectively. Cox regression, meeting the proportional hazards assumption, was employed to obtain hazard ratios (HRs) and 95% confidence intervals (CIs) for the relationships between famine exposure and the risk of overweight, general obesity and abdominal obesity. Furthermore, age, sex, ethnicity, smoking, alcohol consumption, gross family income, physical activity, famine severity and living area at baseline, as well as death status, were adjusted for in the results. WC was further adjusted to examine the independent effects of famine on overweight and general obesity. Similarly, BMI was additionally adjusted to analyze the association of famine exposure with abdominal obesity. The associations of famine exposure with age at onset of overweight, general obesity and abdominal obesity were analyzed using linear regression analyses. The conceptual diagram is presented in Fig. 1. SAS 9.4 (SAS Institute Inc., Cary, NC, USA.) was used to conduct all analyses, and a two-tailed $P \leq 0.05$ indicated statistical significance.

Results

Baseline characteristics

In total, 6957 participants were included in this study. The average age, BMI and WC were 36.94 ± 8.17 years, 22.33 ± 2.90 kg/m^2 and

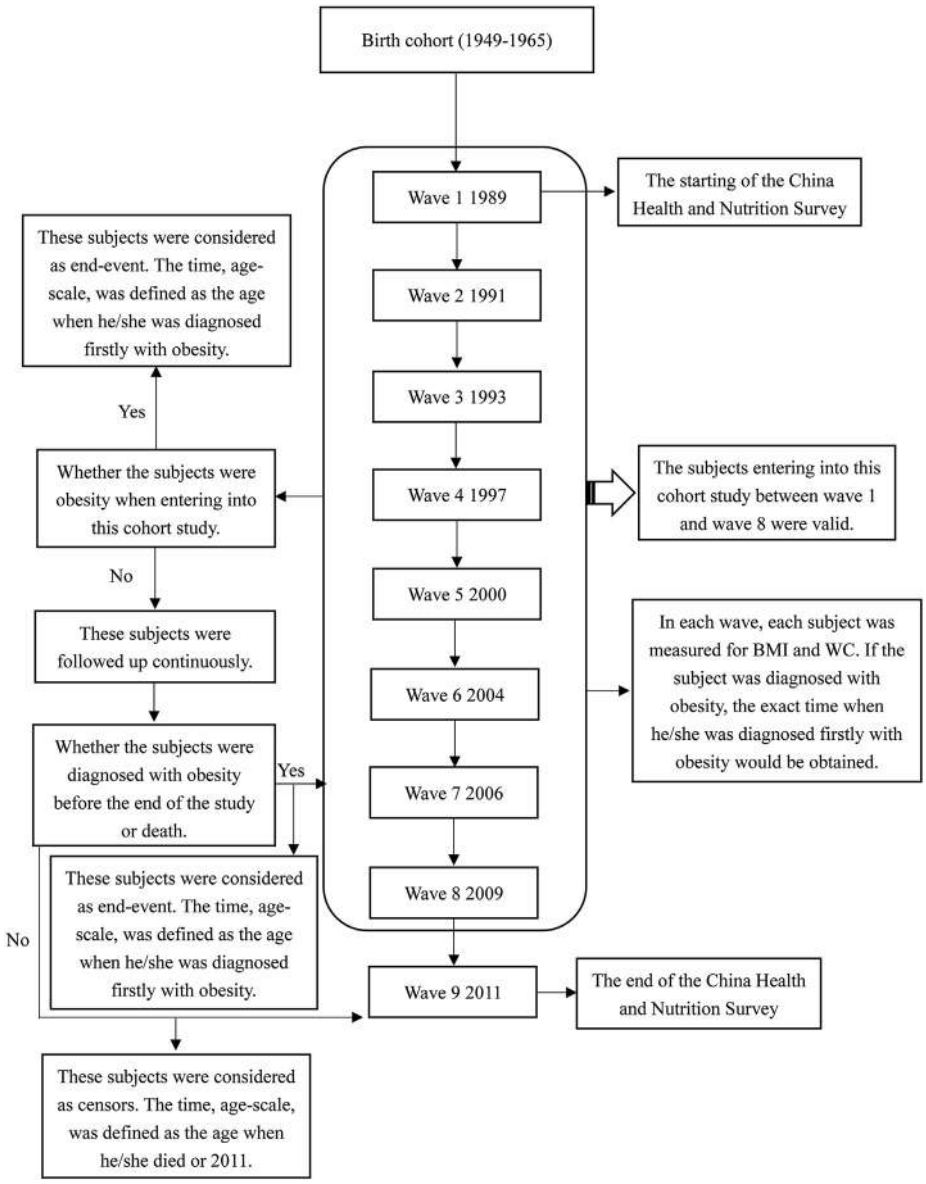


Fig. 1. The conceptual diagram of this study. BMI, body mass index; WC, waist circumference.

80.92 ± 9.62 cm, respectively. The median follow-up period was 43 person-years. Study participant characteristics at baseline across the five famine exposure cohorts are shown in Table 1. Significant differences were observed in age ($P < 0.001$), BMI ($P < 0.001$), WC ($P < 0.001$), gross family income ($P = 0.023$), general obesity ($P < 0.001$) and abdominal obesity ($P < 0.001$) across all five famine exposure cohorts.

Association of exposure to famine in early life with overweight in adulthood

In the total sample, exposure to famine in early life was associated with a lower risk of being overweight (all $P < 0.001$). The results for males and females are consistent with those of the total population. In areas of both 'less severe famine' and 'severe famine', famine exposure decreased the risk of overweight (Table 2). Furthermore, exposure to famine in any period of childhood was linked to a lower risk of overweight in rural and urban areas (see Supplemental Fig. S1).

Association of exposure to famine in early life with general obesity in adulthood

Exposure to famine in mid-childhood and late childhood decreased the risk of general obesity (HR 0.449, 95% CI 0.315–0.640 [$P < 0.001$] and HR 0.571; 95% CI 0.387–0.841 [$P = 0.005$], respectively). There were inverse associations of exposure to famine in mid-childhood with general obesity in males and females (HR 0.485, 95% CI 0.292–0.807 [$P = 0.005$] and HR 0.426, 95% CI 0.256–0.709 [$P = 0.001$], respectively) (Table 2). In rural areas, only famine exposure in mid-childhood decreased the risk of general obesity (HR 0.466, 95% CI 0.284–0.764 [$P = 0.003$]). In urban areas, there were inverse associations of exposure to famine in mid-childhood and late childhood with general obesity (HR 0.445, 95%

Table 1
Participant characteristics at baseline.

Variables	Famine exposure					*P*-Value
	No exposure	Fetal life	Early childhood	Mid-childhood	Late childhood	
Age (years)[a]	30.88 ± 6.47	34.52 ± 6.53	37.47 ± 6.71	40.30 ± 6.57	43.73 ± 6.60	<0.001
BMI (kg/m^2)[a]	22.01 ± 2.84	22.44 ± 2.98	22.36 ± 2.83	22.38 ± 2.73	22.66 ± 3.10	<0.001
WC (cm)[a]	79.75 ± 9.37	80.62 ± 10.09	80.77 ± 9.43	81.62 ± 9.31	82.38 ± 9.91	<0.001
Sex[b]						0.886
Male	1038 (49.45)	424 (47.91)	599 (49.06)	626 (48.27)	723 (49.69)	
Female	1061 (50.55)	461 (52.09)	622 (50.94)	671 (51.73)	732 (50.31)	
Smoking[b]						0.180
No	1201 (61.18)	505 (59.62)	711 (59.85)	770 (60.53)	815 (57.07)	
Yes	762 (38.82)	342 (40.38)	477 (40.15)	502 (39.47)	613 (42.93)	
Alcohol consumption[b]						0.224
No	999 (50.97)	441 (52.07)	597 (50.29)	603 (47.41)	710 (49.72)	
Yes	961 (49.03)	406 (47.93)	590 (49.71)	669 (52.59)	718 (50.28)	
Ethnicity[b]						0.056
Han	1771 (84.37)	749 (84.63)	1055 (86.40)	1075 (82.88)	1201 (82.54)	
Other	328 (15.63)	136 (15.37)	166 (13.60)	222 (17.12)	254 (17.46)	
Physical activity[b]						0.645
No	1991 (94.85)	837 (94.58)	1148 (94.02)	1215 (93.68)	1375 (94.50)	
Yes	108 (5.15)	48 (5.42)	73 (5.98)	82 (6.32)	80 (5.50)	
Gross income family[b]						0.023
Low	1034 (49.26)	429 (48.47)	644 (52.74)	690 (53.20)	774 (53.20)	
High	1065 (50.74)	456 (51.53)	577 (47.26)	607 (46.80)	681 (46.80)	
Famine severity[b]						0.375
Less severe	635 (30.25)	298 (33.67)	392 (32.10)	422 (32.54)	469 (32.23)	
Severe	1464 (69.75)	587 (66.33)	829 (67.90)	875 (67.46)	986 (67.77)	
Region[b]						0.169
Urban	785 (37.40)	369 (41.69)	472 (38.66)	510 (39.32)	590 (40.55)	
Rural	1314 (62.60)	516 (58.31)	749 (61.34)	787 (60.68)	865 (59.45)	
Obesity[b]						<0.001
Normal	1464 (69.75)	561 (63.39)	784 (64.21)	834 (64.30)	861 (59.18)	
Overweight	542 (25.82)	266 (30.06)	375 (30.71)	405 (31.23)	486 (33.40)	
General obesity	93 (4.43)	58 (6.55)	62 (5.08)	58 (4.47)	108 (7.42)	
Abdominal obesity[b]						<0.001
Normal	1334 (78.10)	55 5 (72.93)	778 (74.31)	816 (73.12)	863 (67.11)	
Abdominal obesity	374 (21.90)	206 (27.07)	269 (25.69)	300 (26.88)	423 (32.89)	

[a] These variables were analyzed by analysis of variance. Results presented as mean ± SD.
[b] These variables were analyzed by Chi-square test. Results presented as n (%).

CI 0.266–0.743 [$P = 0.002$] and HR 0.514, 95% CI 0.289–0.915 [$P = 0.024$], respectively) (Supplemental Fig. S2).

Association of exposure to famine in early life with abdominal obesity in adulthood

There were inverse associations of exposure to famine in any period of childhood with abdominal obesity in the total sample (fetal life = HR 0.654, 95% CI 0.579–0.738 [$P < 0.001$]; early childhood = HR 0.489, 95% CI 0.434–0.550 [$P < 0.001$]; mid-childhood = HR 0.374, 95% CI 0.330–0.425 [$P < 0.001$]; and late childhood = HR 0.274, 95% CI 0.239–0.315 [$P < 0.001$]). When the results were stratified by sex and famine severity, they were comparable with those of the total sample (Table 2). Similarly, exposure to famine in any period of childhood decreased the risk of abdominal obesity in rural and urban areas (Supplemental Fig. S3).

Associations of exposure to famine in early life with age at onset of overweight, general, and abdominal obesity

Table 3 shows that exposure to famine in early, mid-childhood and late childhood delayed the age at onset of overweight in the total sample ($P = 0.004$, $\beta = 0.651$; $P < 0.001$, $\beta = 0.972$; and $P < 0.001$, $\beta = 1.594$, respectively). With the exception of the fetal lifetime period, exposure to famine at any other early stage of life delayed the age at onset of general obesity. In the total sample and subgroups, exposure to famine in any period of childhood delayed the age at onset of abdominal obesity.

Discussion

Data from the CHNS were used to examine the associations of exposure to the Chinese Famine with overweight, general and abdominal obesity in the Chinese population. Results showed that exposure to famine was linked to lower risks of overweight and abdominal obesity. However, there was no significant relationship between famine exposure and general obesity. In addition, exposure to famine delayed the age at onset of overweight and abdominal obesity.

Previous investigations were limited to cross-sectional studies. However, the present study was based on a prospective cohort and corrected for the confounding factors of age and famine severity. As a result, unlike previous studies, this study showed that exposure to famine was linked to lower risks of overweight and abdominal obesity, especially in males in the Chinese population. These results are in line with several studies reporting that males who were exposed to famine in the last trimester of gestation or the immediate postnatal period were less likely to experience obesity than those who were not exposed to famine.[26,27] These results may indicate that food intake in childhood considerably impacts obesity in adulthood and that starvation may reduce the risk of obesity.[28] It was well known that undernutrition in early life has far-reaching consequences on health outcomes in later life.[29,30] Famine-related studies using data from the Dutch and Ukrainian famines were conducted to develop the origins hypothesis.[31,32] A previous study by Schulz showed that individuals who had been exposed to famine during late gestation had a lower rate of obesity than those who

Table 2
Associations of exposure to famine with risk of overweight and general and abdominal obesity.

Famine exposure	Overweight			General obesity			Abdominal obesity		
	P-Value	HR	95% CI	*P*-Value	HR	95% CI	*P*-Value	HR	95% CI
Total (N = 6957)									
No exposure	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Fetal life	<0.001	0.737	0.629–0.863	0.561	0.914	0.674–1.238	<0.001	0.654	0.579–0.738
Early childhood	<0.001	0.667	0.568–0.782	0.101	0.766	0.557–1.053	<0.001	0.489	0.434–0.550
Mid-childhood	<0.001	0.534	0.451–0.631	<0.001	0.449	0.315–0.640	<0.001	0.374	0.330–0.425
Late childhood	<0.001	0.455	0.376–0.551	0.005	0.571	0.387–0.841	<0.001	0.274	0.239–0.315
Males (n = 3410)									
No exposure	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Fetal life	0.007	0.742	0.597–0.922	0.483	0.862	0.570–1.304	<0.001	0.622	0.506–0.764
Early childhood	<0.001	0.533	0.422–0.672	0.051	0.626	0.391–1.002	<0.001	0.421	0.341–0.519
Mid-childhood	<0.001	0.440	0.345–0.561	0.005	0.485	0.292–0.807	<0.001	0.316	0.253–0.395
Late childhood	<0.001	0.407	0.308–0.538	0.071	0.589	0.331–1.046	<0.001	0.181	0.140–0.233
Females (n = 3547)									
No exposure	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Fetal life	0.004	0.711	0.564–0.896	0.903	1.029	0.652–1.624	<0.001	0.677	0.581–0.788
Early childhood	0.094	0.827	0.662–1.033	0.925	0.979	0.624–1.535	<0.001	0.542	0.469–0.627
Mid-childhood	<0.001	0.646	0.511–0.817	0.001	0.426	0.256–0.709	<0.001	0.403	0.345–0.470
Late childhood	<0.001	0.499	0.383–0.651	0.067	0.603	0.351–1.037	<0.001	0.335	0.283–0.397
Less severe famine (n = 2216)									
No exposure	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Fetal life	0.003	0.707	0.563–0.887	0.840	1.045	0.682–1.600	<0.001	0.685	0.561–0.835
Early childhood	<0.001	0.664	0.532–0.829	0.280	0.779	0.496–1.225	<0.001	0.485	0.400–0.588
Mid-childhood	<0.001	0.564	0.442–0.721	0.016	0.538	0.324–0.892	<0.001	0.407	0.330–0.501
Late childhood	<0.001	0.424	0.321–0.560	0.023	0.515	0.290–0.911	<0.001	0.306	0.242–0.387
Severe famine (n = 4741)									
No exposure	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Fetal life	0.005	0.724	0.578–0.907	0.255	0.772	0.494–1.206	<0.001	0.634	0.543–0.740
Early childhood	<0.001	0.637	0.505–0.803	0.151	0.714	0.451–1.131	<0.001	0.488	0.420–0.567
Mid-childhood	<0.001	0.492	0.388–0.623	<0.001	0.349	0.208–0.585	<0.001	0.351	0.300–0.412
Late childhood	<0.001	0.454	0.347–0.594	0.051	0.579	0.334–1.002	<0.001	0.255	0.214–0.304

CI, confidence interval; HR, hazard ratio.

Table 3
Associations of exposure to famine with the age at onset of overweight and general and abdominal obesity.

Famine exposure	Overweight		General obesity		Abdominal obesity	
	β	*P*-Value	β	*P*-Value	β	*P*-Value
Total (n = 6957)						
No exposure	Ref	Ref	Ref	Ref	Ref	Ref
Fetal life	0.160	0.484	0.088	0.870	1.488	<0.001
Early childhood	0.651	0.004	1.389	0.011	2.594	<0.001
Mid-childhood	0.972	<0.001	1.978	0.001	3.354	<0.001
Late childhood	1.594	<0.001	2.494	<0.001	4.613	<0.001
Males (n = 3410)						
No exposure	Ref	Ref	Ref	Ref	Ref	Ref
Fetal life	0.507	0.108	0.396	0.577	1.471	<0.001
Early childhood	1.251	<0.001	1.663	0.032	3.306	<0.001
Mid-childhood	1.395	<0.001	1.107	0.175	4.186	<0.001
Late childhood	1.788	<0.001	2.807	0.002	6.001	<0.001
Females (n = 3547)						
No exposure	Ref	Ref	Ref	Ref	Ref	Ref
Fetal life	−0.277	0.409	−0.128	0.876	1.557	<0.001
Early childhood	0.061	0.847	1.356	0.080	2.232	<0.001
Mid-childhood	0.526	0.119	3.302	<0.001	2.951	<0.001
Late childhood	1.342	<0.001	2.571	0.005	3.939	<0.001
Rural (n = 2726)						
No exposure	Ref	Ref	Ref	Ref	Ref	Ref
Fetal life	0.323	0.279	0.596	0.385	1.384	0.001
Early childhood	0.539	0.085	1.304	0.072	2.398	<0.001
Mid-childhood	0.667	0.033	1.715	0.029	2.705	<0.001
Late childhood	1.211	0.001	1.198	0.136	4.526	<0.001
Urban (n = 4231)						
No exposure	Ref	Ref	Ref	Ref	Ref	Ref
Fetal life	0.140	0.684	−0.308	0.706	1.581	<0.001
Early childhood	0.711	0.032	1.567	0.055	2.687	<0.001
Mid-childhood	1.228	0.001	2.224	0.014	3.706	<0.001
Late childhood	1.946	<0.001	4.349	<0.001	4.578	<0.001
Less severe famine (n = 2216)						
No exposure	Ref	Ref	Ref	Ref	Ref	Ref
Fetal life	0.404	0.208	0.639	0.403	0.758	0.054
Early childhood	1.209	<0.001	1.429	0.081	2.479	<0.001
Mid-childhood	1.501	<0.001	1.610	0.070	3.096	<0.001
Late childhood	2.058	<0.001	2.957	0.002	4.027	<0.001
Severe famine (n = 4741)						
No exposure	Ref	Ref	Ref	Ref	Ref	Ref
Fetal life	−0.001	0.997	−0.708	0.350	1.981	<0.001
Early childhood	0.155	0.639	1.610	0.028	2.630	<0.001
Mid-childhood	0.608	0.073	2.688	0.001	3.570	<0.001
Late childhood	1.267	0.001	2.778	0.002	4.981	<0.001

had not been exposed to famine,[33] which was consistent with the conclusions of the current study. Animal experiments found that maternal dietary manipulation could lead to a shortened lifespan and obesity in offspring.[34–36] In this study, the parents of participants in the 'no exposure' famine group inevitably experienced famine prior to reproduction; therefore, their offspring were more likely to develop obesity than their counterparts. In contrast, some studies found that exposure to famine in fetal life was linked to an increased risk of obesity, which differed from the findings of this study.[16,37] Lumey et al. reviewed human studies on the relationship between acute exposure to prenatal famine and adult physical and mental health, which included most famines from the 19th and 20th centuries.[38] This review found that exposure to famine in fetal life promoted the development of obesity. However, the definitions used within each study were inconsistent, which was the main bias affecting the conclusion. Furthermore, whether the confounding effect of age had been corrected for was not mentioned; thus, the results of this review require further examination.

According to the Darwin theory of 'survival of the fittest', the Chinese Famine discussed in this study might have acted as a filter, eliminating comparatively frail individuals and retaining strong individuals.[39] Therefore, famine survivors might be expected to be healthier and might be less likely to develop overweight, general obesity, or abdominal obesity. However, selection bias might exist due to competing risks. The decedent data from the famine were not collected and could not be used to correct for competing risks; thus, National death registry data should be used to examine this bias in future research.

In terms of public health, this study provides evidence for the prevention of obesity. As this study suggests, nutrition status in early life is associated with overweight and abdominal obesity, but not general obesity. Therefore, physicians should pay attention to the nutritional state of pregnant women and the birth weight of offspring in order to prevent obesity in the offspring in adulthood. Obesity interventions are especially important at present because the prevalence of general and abdominal obesity is continuously increasing. In addition, this study indicates that it is necessary to explore the mechanism of how nutrition status in early life impacts the development of obesity in later life.

Strengths and limitations

This is a representative and long-term follow-up cohort study; thus, the results are accurate and fully reflect the true effect of famine exposure on obesity. In this study, age was used as the time scale in Cox regression models; therefore, the confounding effect of

age was considerably reduced. Moreover, the relationship between exposure to famine and the age of onset of obesity further verified the protective effect of famine exposure. Therefore, the results of this study were valid and convincing.

However, some limitations should be noted. First, it is difficult to identify the exact date when the Chinese Famine started and ended, and therefore, there might be misclassification bias in the definition of famine exposure cohorts. Second, because of the lack of specific dietary intakes in the CHNS, obesity-related nutrition was not adjusted for in the results. Third, WC data were not collected before the 1993 wave; thus, the sample size for general obesity differed from that for abdominal obesity. Fourth, some data were not collected in the CHNS, such as mortality data caused by the famine; therefore, there might be unmeasured confounder bias.

Conclusions

Study participants who were exposed to famine were less likely to develop overweight and abdominal obesity compared with individuals who did not experience famine. The protective effects of famine exposure were most obvious for overweight in males and abdominal obesity in both males and females. In addition, exposure to famine postponed the age at onset of overweight and abdominal obesity. However, exposure to famine only had a weak role in the development of general obesity.

Author statements

Acknowledgements

This research used data from China Health and Nutrition Survey (CHNS). We thank the National Institute for Nutrition and Health, China Center for Disease Control and Prevention, Carolina Population Center, the University of North Carolina at Chapel Hill, the NIH, and the NIH Fogarty International Center for the CHNS data collection and analysis files from 1989 to 2015 and future surveys, and the China-Japan Friendship Hospital, Ministry of Health for support for CHNS 2009, Chinese National Human Genome Center at Shanghai since 2009, and Beijing Municipal Center for Disease Prevention and Control since 2011.

Ethical approval

This study was approved by the Institutional Review Board of the National Institute for Nutrition and Food Safety, China Center for Disease Control and Prevention, and the University of North Carolina at Chapel Hill (CHNS-2009–15). All participants provided written informed consent.

Funding

None declared.

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.2021.11.014.

References

1. Tian H, Xie H, Song G, Zhang H, Hu G. Prevalence of overweight and obesity among 2.6 million rural Chinese adults. *Prev Med* 2009;**48**:59–63.
2. Aune D, Sen A, Prasad M, Norat T, Janszky I, Tonstad S, et al. BMI and all cause mortality: systematic review and non-linear dose-response meta-analysis of 230 cohort studies with 3.74 million deaths among 30.3 million participants. *BMJ* 2016;**353**:i2156.
3. Wang Z. Age-dependent decline of association between obesity and mortality: a systematic review and meta-analysis. *Obes Res Clin Pract* 2015;**9**:1–11.
4. Flegal KM, Kit BK, Orpana H, Graubard BI. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *JAMA* 2013;**309**:71–82. the journal of the American Medical Association.
5. Tahergorabi Z, Khazaei M, Moodi M, Chamani E. From obesity to cancer: a review on proposed mechanisms. *Cell Biochem Funct* 2016;**34**(8):533–45.
6. Li Y-C, Li C-L, Qi J-Y, Huang L-N, Shi D, Du S-S, et al. Relationships of dietary histidine and obesity in northern Chinese adults, an internet-based cross-sectional study. *Nutrients* 2016;**8**:420.
7. Ogden CL, Carroll MD, Kit BK, Flegal KM. *Prevalence of obesity in the United States, 2009-2010*. Hyattsville, MD: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics; 2012.
8. 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. *The lancet* 2014;**384**:766–81.
9. Zhang Y, Yang J, Hou W, Arcan C. Obesity trends and associations with types of physical activity and sedentary behavior in US adults: national health and nutrition examination survey, 2007-2016. *Obesity (Silver Spring)* 2021;**29**:240–50.
10. Jaacks LM, Vandevijvere S, Pan A, McGowan CJ, Wallace C, Imamura F, et al. The obesity transition: stages of the global epidemic. *Lancet Diabetes Endocrinol* 2019;**7**:231–40.
11. Chen Y, Peng Q, Yang Y, Zheng S, Wang Y, Lu W. The prevalence and increasing trends of overweight, general obesity, and abdominal obesity among Chinese adults: a repeated cross-sectional study. *BMC Publ Health* 2019;**19**:1293.
12. Wells JC. The programming effects of early growth. *Early Hum Dev* 2007;**83**:743–8.
13. Druet C, Ong KK. Early childhood predictors of adult body composition. *Best Pract Res Clin Endocrinol Metabol* 2008;**22**:489–502.
14. Barker DJ, Godfrey K, Gluckman PD, Harding JE, Owens JA, Robinson JS. Fetal nutrition and cardiovascular disease in adult life. *Lancet* 1993;**341**:938–41.
15. Ky G. *An overview of nutrition sciences*. Beijing: Peoples Medical Publishing House; 2004.
16. Yang Z, Zhao W, Zhang X, Mu R, Zhai Y, Kong L, et al. Impact of famine during pregnancy and infancy on health in adulthood. *Obes Rev* 2008;**9**:95–9.
17. Wang Y, Wang X, Kong Y, Zhang JH, Zeng Q. The Great Chinese Famine leads to shorter and overweight females in Chongqing Chinese population after 50 years. *Obesity* 2010;**18**:588–92.
18. Liu L, Pang Z, Sun J, Xue B, Wang S, Ning F, et al. Exposure to famine in early life and the risk of obesity in adulthood in Qingdao: evidence from the 1959–1961 Chinese famine. *Nutr Metabol Cardiovasc Dis* 2017;**27**:154–60.
19. Ravelli AC, van der Meulen JH, Osmond C, Barker DJ, Bleker OP. Obesity at the age of 50 y in men and women exposed to famine prenatally. *Am J Clin Nutr* 1999;**70**:811–6.
20. Li C, Lumey L. Studies into severe famine in early life and diabetes in adulthood: the need to control for differences in participant age and location. *Diabetologia* 2017;**60**:1359–60.
21. Popkin BM, Du S, Zhai F, Zhang B. Cohort Profile: the China Health and Nutrition Survey—monitoring and understanding socio-economic and health change in China, 1989–2011. *Int J Epidemiol* 2010;**39**:1435–40.
22. Mui P, Hill SE, Thorpe Jr RJ. Overweight and obesity differences across ethnically diverse subgroups of asian American men. *Am J Men's Health* 2018;**12**:1958–65.
23. W. H. O. Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet* 2004;**363**:157–63.
24. Li C, Lumey LH. Exposure to the Chinese famine of 1959-61 in early life and long-term health conditions: a systematic review and meta-analysis. *Int J Epidemiol* 2017;**46**:1157–70.
25. Sun Y, Zhang L, Duan W, Meng X, Jia C. Association between famine exposure in early life and type 2 diabetes mellitus and hyperglycemia in adulthood: Results from the China Health And Retirement Longitudinal Study (CHARLS). *J Diabetes* 2018;**10**(9):724–33.
26. Ravelli G-P, Stein ZA, Susser MW. Obesity in young men after famine exposure in utero and early infancy. *N Engl J Med* 1976;**295**:349–53.
27. Jelliffe DB. *The assessment of the nutritional status of the community*. Geneva Switzerland: World Health Organization; 1966. p. 271 (World Health Organization Monograph Series No 53).
28. Hermanussen M, Danker-Hopfe H, Weber GW. Body weight and the shape of the natural distribution of weight, in very large samples of German, Austrian and Norwegian conscripts. *Int J Obes Relat Metab Disord* 2001;**25**:1550–3.

29. Gluckman PD, Hanson MA, Cooper C, Thornburg KL. Effect of in utero and early-life conditions on adult health and disease. *N Engl J Med* 2008;**359**:61–73.
30. Godfrey KM, Barker DJ. Fetal nutrition and adult disease. *Am J Clin Nutr* 2000;**71**. 1344S-52S.
31. Roseboom TJ, Painter RC, van Abeelen AF, Veenendaal MV, de Rooij SR. Hungry in the womb: what are the consequences? Lessons from the Dutch famine. *Maturitas* 2011;**70**:141–5.
32. Lumey L, Khalangot MD, Vaiserman AM. Association between type 2 diabetes and prenatal exposure to the Ukraine famine of 1932–33: a retrospective cohort study. *The Lancet Diabetes & Endocrinology* 2015;**3**:787–94.
33. Schulz LC. The Dutch Hunger Winter and the developmental origins of health and disease. *Proc Natl Acad Sci U S A* 2010;**107**:16757–8.
34. Ozanne SE, Hales CN. Lifespan: catch-up growth and obesity in male mice. *Nature* 2004;**427**:411–2.
35. Aihie Sayer A, Dunn R, Langley-Evans S, Cooper C. Prenatal exposure to a maternal low protein diet shortens life span in rats. *Gerontology* 2001;**47**:9–14.
36. Breier BH, Vickers MH, Ikenasio BA, Chan KY, Wong WP. Fetal programming of appetite and obesity. *Mol Cell Endocrinol* 2001;**185**:73–9.
37. Chang X, Song P, Wang M, An L. The risks of overweight, obesity and abdominal obesity in middle age after exposure to famine in early life: evidence from the China's 1959–1961 famine. *J Nutr Health Aging* 2018:1–7.
38. Lumey LH, Stein AD, Susser E. Prenatal famine and adult health. *Annu Rev Publ Health* 2011;**32**:237–62.
39. Song S. Does famine have a long-term effect on cohort mortality? Evidence from the 1959–1961 Great Leap Forward Famine in China. *J Biosoc Sci* 2009;**41**:469–91.

ELSEVIER

PUBLIC HEALTH

www.elsevier.com/locate/puhe

Editorial Office

Natalia Camicia
Kate Cunnington
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

RSPH
ROYAL SOCIETY FOR PUBLIC HEALTH
VISION, VOICE AND PRACTICE

Public Health 202 (2022) 58–64

Contents lists available at ScienceDirect

Public Health

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

Original Research

Effects of social participation and physical activity on all-cause mortality among older adults in Norfolk, England: an investigation of the EPIC-Norfolk study

R.S. Fain [a, *], S.A. Hayat [b], R. Luben [b], A.A. Abdul Pari [c], J.L.Y. Yip [d]

[a] Warren Alpert Medical School of Brown University, Providence, RI, USA
[b] Department of Public Health and Primary Care, University of Cambridge, UK
[c] NHS Specialised Commissioning for Mental Health Programme of Care, UK
[d] International Centre for Evidence on Disability, Clinical Research Unit, London School of Hygiene and Tropical Medicine, UK

ARTICLE INFO

Article history:
Received 15 March 2021
Received in revised form
21 October 2021
Accepted 29 October 2021
Available online 8 December 2021

Keywords:
Social capital
Social participation
Mortality
Ageing
England

ABSTRACT

Objectives: There is growing evidence of an association between social participation and improved physical and mental health among older individuals. The aims of this study were to explore the relationship between self-reported participation in groups, clubs, or organizations and all-cause mortality among older adults and examine the role of physical activity as a potential modifier of the health effects of social participation.

Study design: EPIC-Norfolk is a prospective cohort study that recruited 25,639 individuals between the ages of 40 and 79 in Norfolk County, England. This study involved a retrospective analysis of 8623 participants who had returned for the third health check between 2004 and 2011.

Methods: Participants were categorized into those who reported participating socially and those who did not and were stratified by involvement in 0, 1, or 2 or more groups. Cox Proportional Hazards models were constructed to compare all-cause mortality between the groups. Stratum-specific hazard ratios were calculated by physical activity level to assess for effect modification.

Results: Of the participants, 861 (9.98%) died during the follow-up period. After adjustment for confounding, social participation was associated with lower all-cause mortality (HR 0.84, 95% CI 0.73–0.97). Involvement in 2 or more groups was associated with lower all-cause mortality (hazard ratio [HR] 0.83, 95% confidence interval [CI] 0.70–0.97), but the association was not statistically significant for people involved in only 1 group (HR 0.86, 95% CI 0.73–1.03). Physical activity appeared to modify the effect of social participation on mortality.

Conclusions: This study's findings provide evidence of an association between social participation and lower all-cause mortality for older adults. They also suggest that the effect of social participation on health is greater for people who are more physically active. Population-level interventions to facilitate social participation may contribute to improving health and wellbeing among older individuals.

Introduction

In *Bowling Alone: The Collapse and Revival of American Community*, Putnam suggests that forms of social capital such as participation in community groups may improve health through providing tangible assistance, reinforcing health norms, and reducing stress.[1] Critics of social capital as a topic of research point to the term's ambiguity[2] and its reliance on proxy measures for individual exposure.

Psychosocial aspects of health inequalities entered mainstream epidemiology in the 1980s when the Whitehall II study revealed that lower social standing was a significant risk factor for poor health.[3] There is growing evidence of differences in social participation as a driver of health inequalities. A 2019 meta-review of 20 systematic reviews found good evidence in support of the concept that structural social capital, which includes social participation, predicts better mental and physical health, but also highlighted several systematic

* Corresponding author. Warren Alpert Medical School of Brown University, 222 Richmond St, Providence, 02903, RI, USA. Tel.: +1 (419) 564-3301.
E-mail address: robert.fain.14@gmail.com (R.S. Fain).

https://doi.org/10.1016/j.puhe.2021.10.017
0033-3506/

reviews that showed non-significant or negative associations between social capital and health outcomes.[4] Heterogeneity of results likely owes to differences in demographic makeup and social dynamics among study populations. A 2017 systematic review of 44 studies by Villalonga-Olives and Kawachi explored contexts in which social capital was associated with harmful health behaviors and worse mental and physical health outcomes, such as increased levels of smoking and drinking among Japanese youth with higher participation in extracurricular activities.[5]

Another factor that varies across studies is how researchers address confounding and interaction. Physical activity, for example, is well supported in the literature as protective against mortality.[6] Research also suggests that physical activity and social participation share a reciprocal relationship, positively reinforcing one another.[7–13] Adjustment for physical activity as a confounder could underestimate the effect of social participation; however, researchers should endeavor to account for the role of physical activity and its potential influence on the health effects of social participation.

The primary aim of the present study was to explore the relationship between social participation and all-cause mortality in the United Kingdom, using longitudinal models. The secondary aim was to assess the potential role of physical activity as a modifier of the effects of social participation. Although many studies have examined associations between social participation and long-term health outcomes, few studies have researched potential interrelated effects of social participation and physical activity on mortality.

If social participation has the potential to reduce mortality and chronic disease burden, investment in neighborhood-level resources that promote social capital could serve as an effective method of reducing health inequalities in communities. Such interventions could have an important effect on older people, who are at higher risk of social isolation or disability.

Methods

Study population and setting

The European Prospective Investigation of Cancer (EPIC) is a 10-country cohort study designed to examine social and environmental determinants of health, with detailed methods published elsewhere.[14] This analysis focuses on the EPIC-Norfolk cohort, one of the UK sites of the study. As nearly all residents in the United Kingdom are registered with a general practitioner, general practice lists effectively serve as population registers. Between 1993 and 1997, 25,639 participants aged 40–79 were recruited via general practices to receive examinations over the follow-up period.

The present study uses data from the Third Health Check (3HC), conducted between 2006 and 2011. In the 3HC, 18,380 men and women between ages 48 and 92 were contacted, of whom 8623 (47%) were examined.[15] All participants completed a detailed self-administered health and lifestyle questionnaire and attended a local clinic for a physical examination.

EPIC-Norfolk was carried out following the principles of the Declaration of Helsinki and the Research Governance Framework for Health and Social Care. The study was approved by the Norfolk Local Research Ethics Committee (05/Q0101/191) and East Norfolk and Waveney NHS Research Governance Committee (2005EC07L). All participants gave written, informed consent.

A health examination was carried out by trained nurses following standard operating protocols. Height and weight were measured with participants dressed in light clothing and shoes removed. A stadiometer was used to measure height to the nearest 0.1 cm (cm), and the Tanita body composition analyzer model TBF 300s (Chasmors Ltd, London) was used to measure weight to the nearest 100 g.

Participants were also sent two questionnaires. The first questionnaire covered demographic, lifestyle, health, and wellness factors, and the second included social factors.[16,17] Deaths, causes, and dates through the follow-up end date of March 31, 2016, were obtained through death certificates obtained using linkage to NHS Digital. Participants in the dataset remained anonymous.

Explanatory variables

Continuous variables included age and body mass index (BMI). Townsend deprivation index scores, measures of the areas' relative material deprivation based on unemployment, non-car ownership, non-home ownership, and household overcrowding, were grouped into quartiles to account for the distribution's right skew.[18] Smoking was recorded as current, former or never, and alcohol consumption was recorded as units consumed in the past week using the amount of beer, wine, spirits and fortified wine consumed to estimate units. Educational attainment was classified into less than O-level, up to and including O-Level, up to and including A-Level, university degree, or postgraduate qualifications according to the highest qualification achieved. Social class was recorded as professional, manager, skilled non-manual, skilled manual, semi-skilled and non-skilled, using the Registrar General's occupation-based classification system. Participants described their physical activity as one of four levels: inactive, moderately inactive, moderately active, or active.

The dataset presented several potential measures for individual social capital. Group membership was chosen as the explanatory variable because it served as a generalizable individual-level measure of a community-level resource.[19]

Social participation

The EPIC-Norfolk Health Questionnaire asked participants if they regularly joined in the activities of a list of organizations.[20] Social participation was classified based on self-reported responses across thirteen social activities: (1) political parties, (2) trade unions, (3) environmental groups, (4) parent–teacher associations, (5) residents' associations, (6) classes, (7) charity groups, (8) groups for elderly people, (9) youth groups, (10) women's groups, (11) social clubs, (12) sports clubs, and (13) other group or organization. A final option was available stating, 'No, I don't regularly join in any of the activities of these organizations'.[20] Data on social participation were dichotomized comparing participants who reported participation in at least one group and did not check the final option, to those who did not report group participation and/or selected the final option.[20] An additional variable was generated that stratified participants who participated socially into those who were involved in one group and those who were involved in two or more groups.

Physical activity was treated as a potential effect modifier between social participation and mortality. This approach was informed by the Social Ecological Model[21–23] and Barton and Grant's 'health map'.[24]

Statistical analysis

The time-to-event analysis explored the effects of social participation on long-term health outcomes. Data were analyzed using STATA 14 and 15 (Statacorp, Texas). Means, proportions, and standard deviations were calculated.

Associations between potential confounders and explanatory variables of interest were examined using Analysis of variance and chi-square tests. Associations between social participation, number of groups joined, other variables of interest, and all-cause mortality were measured via Cox Proportional Hazards regression. Potential

confounders were added in a stepwise fashion, and goodness of fit was determined using Likelihood Ratio Tests comparing models with and without the additional variables. Adjustment for confounding by Townsend deprivation quartile or education level did not improve the models' goodness of fit, and these variables were not included in the fully adjusted models. Final models, which adjusted for age, gender, smoking status, and BMI, were constructed for both social participation and the number of groups joined. Physical activity, gender, education level, and Townsend deprivation quartile were tested for effect modification using Cox Proportional Hazards models with interaction terms.

Results

There were 8623 participants in the sample, with women comprising 55% (n = 4762) of the study population. Over 99% of participants were of European descent.[25] Although those who attended 3HC tended to be younger and have a higher socioeconomic position (SEP) compared to those who did not, the sample still represented a wide range of socioeconomic characteristics.[14] The average follow-up period for participants was 7.3 years. Of the 8623 participants, 861 (9.98%) died during the follow-up period [Table 1].

Women, non-smokers, people with lower BMI, higher levels of education, those who lived in wealthier neighborhoods and worked in non-manual occupations were more likely to participate socially [Table 2].

After adjusting for age, gender, smoking status, and BMI, the hazard ratio (HR) for all-cause mortality among those who participated socially was 0.84 (95% confidence interval [CI]: 0.73–0.97) (see Table 3). The hazard ratio of all-cause mortality among participants who were involved in one group was 0.86 (95% CI: 0.73–1.03), and 0.83 (95% CI: 0.70–0.97) among those who were involved in two or more groups [Table 4].

Physical activity appeared to modify the effect of social participation on all-cause mortality. The association between social participation and lower mortality was strongest among people who reported being physically active (HR 0.66; 95% CI 0.39–1.11). The effect size decreased for people who reported being moderately active (HR 1.00; 95% CI 0.65–1.54) or moderately inactive (HR 0.84; 95% CI 0.58–1.21). Sex, education level, and Townsend deprivation index score did not modify the effect of social participation on all-cause mortality.

Table 1
EPIC-Norfolk Cohort sample baseline characteristics at the 3rd Health Check (n = 8623).

Variable	Level	(n = 8623)	
		Mean	(SD)
Age (years)		68.7	(8,1)
Body mass index (kg/m^2)		26.8	(4.3)
		%	(n)
Sex	Female	55.2	4762
	Male	44.8	3861
Smoking status	Current	4.4	372
	Former	46.0	3909
	Never	49.6	4220
Townsend deprivation quartile	1 (least deprived)	25.2	2163
	2	24.2	2083
	3	25.6	2202
	4 (most deprived)	25.0	2153
Social class	Professional	8.8	750
	Manager	41.1	3511
	Skilled non-manual	16.1	1376
	Skilled manual	20.5	1753
	Semi-skilled	11.2	954
	Non-skilled	2.3	199
Education level	None/Primary only	26.3	2269
	O-level	11.9	1026
	A-level	44.2	3810
	Degree	17.6	1516
Participates in social groups?	No	37.7	3254
	Yes	62.3	5369
Number of groups joined	0	37.7	3254
	1	26.0	2244
	2 or more	36.2	3125
Self-rated health	Excellent	7.6	638
	Very good	35.1	2959
	Good	41.4	3493
	Fair	14.3	1209
	Poor	1.7	139
Physical activity	Inactive	37.3	3170
	Moderately inactive	29.0	2467
	Moderately active	17.8	1509
	Active	15.9	1355
Alive or deceased at end of follow-up period	Alive	90.0	7762
	Deceased	10.0	861

Discussion

Main findings of this study

Our findings suggest that a relationship may exist between social participation and lower all-cause mortality. While the impacts of social participation on health are unlikely to be universal across different settings, in this cohort, participants who reported that they participated socially at baseline had lower hazard rates of mortality compared to those who did not report that they participated socially. Associations remained after adjustment for confounding factors, including age, BMI, and smoking.

There was some evidence that a greater quantity of social participation was associated with lower all-cause mortality: participants who reported involvement in two or more groups had a lower rate of mortality than participants who were members of only one group or no groups. This finding was inconclusive; however, as the sample size was not large enough to determine whether mortality differed between participants involved in one group and participants who were not members of any group. It is not immediately clear why participants who were members of two or more social groups exhibited lower mortality rates than participants who were members of only one group. Higher levels of social participation may improve individuals' abilities to access community resources and create more opportunities for positive health behaviors. It may increase the amount of time individuals spend outside the home and promote more active lifestyles. Equally, the quantity of social participation could be a proxy for an unmeasured variable that represents an underlying difference between people who join multiple groups and people who join just one or no groups. Additional research will be needed to further examine potential relationships between the quantity of social capital and health outcomes.

In this sample, associations between social participation and lower mortality were strongest among participants who were more physically active. This observation supports the notion that social participation and physical activity may mutually reinforce one another. Forms of social participation that promote physical activity may influence health to a greater degree than forms that do not motivate more active lifestyles. Future studies may investigate how forms of social participation may differ among people with different levels of physical activity.

What is already known on this topic

This study contributes to the growing body of literature on the potential long-term health effects of social capital and social participation. A systematic review of 60 studies by Uphoff et al.

Table 2
Comparison of sociodemographic characteristics of EPIC-Norfolk participants based on social participation status and number of groups joined.

N = 8623	Social Participation Measure											
	Social Participation					Number of Groups						
	No		Yes		*P*-value	0		1		**2+**		*P*-value
	(n = 3254)		(n = 5369)			(n = 3254)		(n = 2244)		(n = 2473)		
	mean	(SD)	mean	(SD)		mean	(SD)	mean	(SD)	mean	(SD)	
Age	68.5	(8.3)	68.8	(8.0)	0.10	68.5	(8.3)	68.2	(8.1)	69.3	(7.8)	<0.01
Body Mass Index (kg/m^2)	27.1	(4.4)	26.7	(4.3)	<0.01	27.1	(4.4)	26.7	(4.3)	26.7	(4.3)	<0.01
	%	(n)	%	(n)		%	(n)	%	(n)	%	(n)	
Smoking Status					<0.01							<0.01
Current	55.7	(207)	44.3	(165)		55.7	(207)	23.9	(89)	20.4	(76)	
Former	39.5	(1544)	60.5	(2365)		39.5	(1544)	26.8	(1047)	33.7	(1318)	
Never	32.7	(1381)	67.3	(2839)		32.7	(1381)	26.3	(1108)	41.0	(1731)	
Townsend Deprivation Quartile					<0.01							<0.01
1 (least deprived)	35.1	(759)	64.9	(1404)		35.1	(759)	28.2	(610)	36.7	(794)	
2	36.1	(751)	64.0	(1332)		36.1	(751)	26.9	(560)	37.1	(772)	
3	40.0	(880)	60.0	(1322)		40.0	(880)	24.7	(626)	35.3	(778)	
4 (most deprived)	40.0	(859)	60.0	(1294)		40.0	(859)	24.3	(168)	35.8	(770)	
Education level					<0.01							<0.01
None/Primary only	51.3	(1165)	48.7	(1104)		51.3	(1165)	26.3	(597)	22.3	(507)	
O-levels	38.2	(392)	61.8	(634)		38.2	(392)	29.1	(299)	32.7	(335)	
A-levels	35.4	(1350)	64.6	(2460)		35.4	(1350)	26.4	(1004)	38.2	(1456)	
Degree	22.8	(345)	77.2	(1171)		22.8	(345)	22.7	(344)	54.6	(827)	
Gender					<0.01							<0.01
Men	42.6	(1643)	57.5	(2218)		42.6	(1643)	28.0	(1080)	29.5	(1138)	
Women	33.8	(1611)	66.2	(3151)		33.8	(1611)	24.4	(1164)	41.7	(1987)	
Physical Activity					<0.01							<0.01
Inactive	48.2	(1529)	51.8	(1641)		48.2	(1529)	24.9	(789)	26.9	(852)	
Moderately inactive	30.6	(756)	69.4	(1711)		30.6	(756)	25.3	(625)	44.0	(1086)	
Moderately active	30.6	(462)	69.4	(1047)		30.6	(462)	28.2	(426)	41.2	(621)	
Active	28.4	(385)	71.6	(970)		28.4	(385)	29.8	(404)	41.8	(566)	

found evidence of social capital as a buffer against negative health effects due to socioeconomic inequality.[27] A review of 14 studies by Choi et al., in contrast, reported no evidence of an association between several forms of social capital and all-cause mortality, CVD or cancer.[28] Both reviews noted a lack of high-quality studies, however.

Comparing weighted survey data across 39 US states, Kawachi et al. noted an association between group membership and lower all-cause and coronary heart disease mortality.[29] Controlling for demographic factors only, a 2016 study found social participation to be associated with reduced mortality risk.[30] In addition, several European studies found evidence of civic and social participation protecting against CVD risk factors.[31–33] Conversely, a longitudinal study on social participation and coronary heart disease found no association after adjusting for physical activity and self-rated health.[34]

Several studies have examined the relationships between quantity of group membership and positive health behaviors[35] and outcomes. A 2016 matched cohort study by Steffens et al. found that, among older individuals in England, those who remained active in two or more social groups after retirement had a lower risk of death over the next 6 years compared to those who were active in one or no groups after retirement.[36] Other studies have found associations between multiple community groups membership and subjective well-being, as well as improved mental health.[37–39]

In a 2012 study, Kanamori et al. compared the incidence of functional disability among four groups from a cohort of older people in Japan: active participants in sports clubs, passive participants in sports clubs, people who exercise alone, and sedentary individuals.[40] The authors found that, while active participants had the lowest incidence of disability, there was no significant difference in the incidence of disability between passive participants and people who exercised alone after adjusting for confounders. Such findings suggest that involvement in a sports club is beneficial for one's health not only because it promotes physical activity but also because it facilitates social participation.

Future studies may help determine whether a causal relationship exists between social participation and long-term health outcomes. Kawachi and other epidemiologists theorize that forms of social capital like social participation influence health through three pathways: (1) 'social contagion', the spread of norms and health behaviors; (2) 'informal social control', the ability of a community to maintain order and sanction deviant behavior; and (3) 'collective efficacy', a group's mobilizing potential for taking collective action.[41] The first pathway may help explain how participants in the EPIC-Norfolk benefited from social participation: regular contact with peers through social groups may promote positive health behaviors such as physical activity and decrease social isolation.

What this study adds

The present study had the advantage of analyzing a large prospective cohort followed for over a decade, as well as obtaining data on numerous exposure measures. The longitudinal design reduced bias from reverse causality. The large sample size in this study reduced the likelihood that results were due to random error.

Although many studies on social capital and health have adjusted for physical activity as a confounder or did not include physical activity in the analysis, few studies have explored the potential interaction between social capital and physical activity. Our findings suggest that physical activity may modify the effect of social participation on health outcomes, although further statistical analysis such as employing causal inference models would provide

Table 3
Results from univariable analysis exploring the associations between explanatory variables and all-cause mortality.

N = 8623	Hazard Ratio	(SE)	*P*-value	95% CI
Age	1.13	(0.005)	<0.01	1.12–1.14
Body Mass Index (kg/m^2)	1.01	(0.008)	0.06	1.00–1.03
Smoking Status				
Current	–			
Former	1.13	(0.18)	0.43	0.83–1.56
Never	0.70	(0.12)	0.03	0.51–0.97
Townsend Deprivation Quartile				
1 (Least deprived)	–			
2	1.01	(0.10)	0.91	0.84–1.23
3	1.05	(0.10)	0.61	0.87–1.26
4 (Most deprived)	0.92	(0.09)	0.40	0.76–1.12
Education Level				
No qualifications	–			
O-level	0.58	(0.07)	<0.01	0.45–0.75
A-level	0.76	(0.06)	<0.01	0.65–0.89
Degree	0.72	(0.07)	<0.01	0.58–0.88
Gender				
Men	–			
Women	0.57	(0.04)	<0.01	0.50–0.66
Social Class				
Professional	–			
Manager	1.29	(0.18)	0.06	0.99–1.70
Skilled non-manual	1.33	(0.20)	0.06	0.99–1.79
Skilled manual	1.05	(0.16)	0.73	0.78–1.42
Semi-skilled	1.04	(0.17)	0.82	0.75–1.44
Non-skilled	1.35	(0.33)	0.21	0.84–2.17
Physical Activity Status				
Inactive	–			
Moderately inactive	0.45	(0.04)	<0.001	0.38–0.54
Moderately active	0.49	(0.05)	<0.001	0.40–0.60
Active	0.43	(0.03)	<0.01	0.37–0.49
Self-Rated Health				
Good/fair/poor	–			
Excellent/very good	0.42	(0.03)	<0.01	0.35–0.49
Participates in Social Groups				
No	–			
Yes	0.79	(0.05)	<0.01	0.69–0.91
Number of Groups Joined				
0	–			
1	0.78	(0.07)	<0.01	0.65–0.92
2 or more	0.81	(0.06)	0.01	0.69–0.94
Participates in a sports group				
No	–			
Yes	0.55	(0.05)	<0.01	0.45–0.66

CI, confidence interval.

stronger support for a reciprocal relationship between physical activity and social participation.

This study also adds to the emerging literature on multiple group membership and health outcomes among older individuals, supporting the findings of other studies that have shown associations between involvement in multiple social groups and lower mortality.

Our findings support the concept that social opportunities are important for older individuals' health and well-being. The UK's National Health Service (NHS) has developed guidelines for 'social prescribing', where patients with multiple morbidities are referred to local community groups and agencies for practical assistance and emotional support.[42]

As community-level resources are important ingredients for social participation, interventions to increase and improve access to it may benefit from a population-level approach. Local authorities may consider expanding infrastructure that facilitates community members' access to existing social groups, clubs, and organizations, such as subsidized or free transportation or meeting spaces.

Limitations of this study

This study was limited in its assessment of the impact of physical activity on health outcomes both independently and as a potential modifier of the effects of social participation. As social participation and physical activity were measured at the same time in this cohort, it was not possible to empirically evaluate whether a causal relationship existed between these two variables.

Another limitation of this study was the lack of stratification between individuals who participated in exercise-related social groups, such as sports clubs, and those who only participated in non-exercise groups. As participants were randomly recruited from general practices, only individuals who were healthy or cognitively capable enough were likely to attend baseline and follow-up examinations, leading to potential healthy volunteer bias.[14] Loss to follow-up over the different stages of EPIC-Norfolk was an additional limitation. If participants who had poorer health and/or demonstrated lower social participation were more likely to drop out of the study, their absence in the analysis could lead to an underestimation of effect; however, all participants were flagged for mortality using national databases. Similar trends in missing values for the variables smoking status, physical activity, and self-rated health may have impacted the observed relationships between smoking, social capital, and health behaviors. Recall bias may also have occurred when participants were asked to complete mailed questionnaires, particularly when reporting past behaviors, experiences, or emotions.

Furthermore, the choice of group membership as the explanatory variable of interest had some potential for bias. The question of whether participants regularly joined social group activities was open to interpretation: 'regularly' could mean once a year for some participants and once a week for others. The dataset did not contain a variable for frequency of social participation, so group

Table 4
Cox proportional hazards models: rates of all-cause mortality in people who participate socially in groups compared to people who do not participate socially, and further stratifying results by involvement in 1 or 2+ groups, with calculated likelihood ratios (LR) comparing goodness of fit to previous model.

	Level	Age and gender adjusted			Age, gender and area deprivation adjusted			Age, gender and education level adjusted			Age, gender, smoking status and BMI adjusted		
		HR	(95%CI)	LR	HR	(95%CI)	LR	HR	(95%CI)	LR	HR	(95%CI)	LR
Hazard ratio of all-cause mortality in people who participate socially at baseline (N = 8623)		0.81	(0.71, 0.93)	0.04	0.81	(0.71, 0.93)	0.19	0.81	(0.71, 0.93)	0.93	0.84	(0.73–0.97)	0.03
Hazard ratio of all-cause mortality in people who report participating in 1, or 2+ groups compared to people who do not participate in groups (N = 8623)	1 group 2 or more groups	0.83 0.80	(0.56, 0.98) (0.68, 0.93)	<0.01	0.84 0.80	(0.70–0.99) (0.68–0.93)	0.19	0.83 0.80	(0.70–0.99) (0.68–0.93)	0.94	0.86 0.83	(0.73–1.03) (0.70–0.97)	0.03

CI, confidence interval; HR, hazard ratio.

membership was used as an imperfect alternative to capture participants' quantity of social participation.

Residual confounding may also impact the observed effects of social participation on health. As participants recruited were middle-aged at baseline, the study did not obtain data on many early-life experiences that affect both individuals' social capital and their health outcomes, such as adverse childhood experiences or childhood illnesses. Similarly, the study did not obtain data on disabilities or functional limitations, which are more prevalent in older populations and may limit individuals' opportunities for social participation. As we did not incorporate these variables in the analysis, the effect size may be overestimated.

Conclusion

In summary, this epidemiological study on the effects of social participation on all-cause mortality is encapsulated within a large, ongoing cohort study of a mostly white British population. Those who participated in a social group, club, or organization were more likely to be women, more highly educated, and in non-manual professions. This study provides evidence in support of an association between social participation and reduced all-cause mortality and suggests that membership in two or more groups may demonstrate an even stronger relationship with lower mortality. Associations between social participation and lower mortality were strongest among people who were physically active.

Author statements

Ethical approval

Ethical approval was not required for this study as no new data were collected.

Funding

No funding was received to support the writing of this paper.

Competing interests

The authors declare that they have no conflict of interest.

References

1. Putnam RD. *Bowling alone: the Collapse and revival of American community*. New York: Simon & Schuster; 2001.
2. Wakefield SEL, Poland B. Family, friend of foe? Critical reflections on the relevance and role of social capital in health promotion and community development. *Soc Sci Med* 2005;**60**:2819–32.
3. Marmot M. *Status syndrome: how social standing affects our health and longevity*. New York: Holt Paperbacks; 2004.
4. Ehsan A, Klass HS, Bastianen A, Spini D. Social capital and health: a systematic review of systematic reviews. *SSM – Population Health* 2019;**8**:100425.
5. Villalonga-Olives E, Kawachi I. The dark side of social capital: a systematic review of the negative health effects of social capital. *Soc Sci Med* 2017;**194**:105–27.
6. Lear SA, Hu W, Rangarajan S, Gasevic D, Leong D, Iqbal R, et al. The effect of physical activity on mortality and cardiovascular disease in 130,000 people from 17 high-income, middle-income, and low-income countries: the PURE study. *Lancet* 2017;**390**:2643–54.
7. Corbett DB, Rejeski WJ, Tudor-Locke C, Glynn NW, Kritchevsky SB, McDermott MM, et al. Social participation modifies the effect of a structured physical activity program on major mobility disability among older adults: results from the LIFE study. *J Gerontol B Psychol Sci Soc Sci* 2018;**73**(8):1501–13.
8. Stevens M, Cruwys T. Membership in sport or exercise groups predicts sustained physical activity and longevity in older adults compared to physically active matched controls. *Ann Behav Med* 2020;**54**(8):557–66.
9. Kawachi I, Kennedy BP, Glass R. Social capital and self-rated health: a contextual analysis. *Am J Publ Health* 1999;**89**:1187–93.
10. Sundquist K, Yang M. Linking social capital and self-rated health: a multilevel analysis of 11,175 men and women in Sweden. *Health Place* 2007;**13**(2):324–34.
11. Lindström M, Moghaddassi M, Merlo J. Social capital and leisure time physical activity: a population based multilevel analysis in Malmö, Sweden. *J Epidemiol Community Health* 2003;**57**:23–8.
12. Ueshima K, Fujiwara T, Takao S, Suzuki E, Iwase T, Doi H, et al. Does social capital promote physical activity? A population-based study in Japan. *PLoS One* 2010;**5**(8):e12135.
13. Chen WL, Zhang CG, Cui ZY, Wang JY, Zhao J, Wang JW, et al. The impact of social capital on physical activity and nutrition in China: the mediating effect of health literacy. *BMC Publ Health* 2019;**19**:1713.
14. Hayat SA, Luben R, Keevil VL, Moore S, Dalzell N, Bhaniani A, et al. Cohort Profile: a prospective cohort study of objective physical and cognitive capability and visual health in an ageing population of men and women in Norfolk (EPIC-Norfolk 3). *Int J Epidemiol* 2013:1–10.
15. EPIC-Norfolk. Methods: timetable. University of Cambridge. N.D. Web. Accessed 8 September 2017. http://www.srl.cam.ac.uk/epic/about/timetable.shtml.
16. EPIC-Norfolk. Institute of Public Health, University of Cambridge. 3HC Confidential. Health Questionnaire.
17. EPIC-Norfolk. Institute of Public Health, University of Cambridge. 3HC Confidential. Health and Life Experiences Questionnaire 2.
18. National Centre for Research Methods. Townsend deprivation index. Geographical Referencing Learning Resources. Web. Accessed 1 October 2021. https://www.restore.ac.uk/geo-refer/36229dtuks00y19810000.php.
19. Nahapiet J, Sumantra G. Social capital, intellectual capital, and the organizational advantage. *Acad Manag Rev* 1998;**23**(2):242.
20. EPIC-Norfolk. 3HC Confidential. Health questionnaire.
21. Bronfenbrenner U. Ecological systems theory. In: Greenwich Vasta R, editor. *Six theories of child development: revised formulations and current issues*, vol. 6. Connecticut: JAI Press; 1989.
22. McLeroy, 1988 24McLeroy KR. An ecological perspective on health promotion programs. *Health Educ Behav* 1988;**15**(4).
23. Ohri-Vachaspati P, DeLia D, DeWeeze RS, Crespo NC, Todd M, Yedidia MJ. The relative contribution of layers of the Social Ecological Model to childhood obesity. *Publ Health Nutr* 2015;**18**(11):2055–66.
24. Barton H, Grant M. A health map for the local human habitat. *J Roy Soc Promot Health* 2006;**126**(6):252–3.
25. Khawaja AP, Chan MPY, Hayat S, Broadway DC, Luben R, Garway-Heath DF, et al. The EPIC-Norfolk Eye Study: rationale, methods and a cross-sectional analysis of visual impairment in a population-based cohort. *BMJ Open* 2013;**3**:e002684.
27. Uphoff EP, Pickett KE, Cabieses B, Small N, Wright J. A systematic review of the relationships between social capital and socioeconomic inequalities in health: a contribution to understanding the psychosocial pathways of health inequalities. *Int J Equity Health* 2013;**12**:54.
28. Choi M, Mesa-Frias M, Nuesch E, Hargreaves J, Prieto-Merino D, Bowling A, et al. Social capital, mortality, cardiovascular disease and cancer: a systematic review of prospective studies. *Int J Epidemiol* 2014:1–26.
29. Kawachi I, Kennedy BP, Lochner K, Prothrow-Stith D. Social capital, income inequality and mortality. *Am J Publ Health* 1997;**87**:1491–8.
30. Steffens NK, Cruwys T, Haslam C, Jetten J, Alexander Haslam S. Social group memberships in retirement are associated with reduced risk of premature death: evidence from a longitudinal cohort study. *BMJ Open* 2016;**6**:1–9.
31. Ellaway A, Macintyre S. Is social participation associated with cardiovascular disease risk factors? *Soc Sci Med* 2007;**64**:1384–91.
32. Sundquist K, Lindström M, Malmström M, Johansson SE, Sundquist J. Social participation and coronary heart disease: a follow-up study of 6900 women and men in Sweden. *Soc Sci Med* 2004;**58**:615–22.
33. Sundquist J, Johansson SE, Yang M, Sunquist K. Low linking social capital as a predictor of coronary heart disease in Sweden: a cohort study of 2.8 million people. *Soc Sci Med* 2006;**62**:954–63.
34. Floud S, Balkwill A, Canoy D, Reeves GK, Green J, Beral V, et al. Social participation and coronary heart disease risk in a large prospective study of UK women. *European Journal of Preventive Cardiology* 2016;**23**(9):995–1002.
35. Sani F, Madhok V, Norbury M, Dugard P, Wakefield JRH. Greater number of group identifications is associated with healthier behaviour: evidence from a Scottish community sample. *Br J Health Psychol* 2014;**20**(3):466–81.
36. Steffens NK, Cruwys T, Haslam C, Jetten J, Haslam SA. Social group memberships in retirement are associated with reduced risk of premature death: evidence from a longitudinal cohort study. *BMJ Open* 2016;**6**:e010164.
37. Fancourt D, Steptoe A. Community group membership and multidimensional subjective well-being in older age. *Ageing and health* 2018;**72**:376–82.
38. Cruwys T, Dingle GA, Haslam SA, Haslam C, Jetten J, Morton TA. Social group memberships protect against future depression, alleviate depression symptoms and prevent depression relapse. *Soc Sci Med* 2013;**98**:179–86.
39. Jetten J, Branscombe NR, Haslam SA, Haslam C, Cruwys T, Jones JM, et al. Having a lot of a good thing: multiple important group memberships as a source of self-esteem. *PLoS One* 2015;**10**:e0124609.

40. Kanamori S, Kai Y, Kondo K, Hirai H, Ichida Y, Suzuki K, et al. Participation in sports organizations and the prevention of functional disability in older Japanese: the AGES cohort study. *PLoS One* 2012;**7**(11):e51061.
41. Kawachi I, Kennedy BP, Lochner K, Prothrow-Stith D. Social capital, income inequality and mortality. *Am J Publ Health* 1997;**87**:1491–8 [It is important to note that these group-level characteristics are not inherently beneficial to health. Social capital may be thought of as a neutral resource, whose positive, or possibly negative, effect on one's health depends on both the broader social environmental context and one's position in relation to the rest of society].
42. NHS England. *Social prescribing and community-based support: summary guide*. 2019.

Public Health 202 (2022) 26–31

Contents lists available at ScienceDirect

Public Health

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

Original Research

Gender differences in specific trends of COPD mortality in Croatia

D. Jovičić Burić [a, *], M. Erceg [a], N. Antoljak [a, b]

[a] Croatian Institute of Public Health, Zagreb, Croatia
[b] The University of Zagreb, School of Medicine, The 'Andrija Štampar' School of Public Health, Zagreb, Croatia

ARTICLE INFO

Article history:
Received 29 June 2021
Received in revised form
17 October 2021
Accepted 29 October 2021
Available online 5 December 2021

Keywords:
Mortality
Trends
COPD
Pulmonary disease
Croatia
Epidemiology

ABSTRACT

Objectives: Chronic obstructive pulmonary disease (COPD) is one of the most common lower respiratory chronic diseases. The aim of this study was to analyze the COPD mortality trends in Croatia for the period 2010–2019 and to identify possible changes and differences by age group and gender.
Study design and methods: In data analysis were included COPD death cases for the period 2010–2019 defined as ICD-10 code J44.0 – J44.9. Mortality data were obtained from the Croatian Institute of Public Health based on death certificates. To model temporal changes in mortality rates joinpoint regression analysis was carried out.
Results: The number of COPD deaths increased in men from 878 in 2010 to 1083 in 2019 and in women from 520 in 2010 to 737 in 2019. Over the 10-year period, there was a stable age-standardized COPD mortality rate among men and statistically significant increasing age-standardized COPD mortality rate among women at the national level.
Conclusions: The findings show a narrowing of the gender gap of COPD mortality. Observed higher COPD mortality rates with age in both men and women confirm previous data and imply that the number of COPD deaths will continue to increase in the future. The healthcare system should focus on the improvement of the quality of care and investment in health promotion and prevention programs aimed at reducing risk factors for COPD, especially tobacco smoking, as well as raising awareness and knowledge about COPD as a chronic disease.

Introduction

Chronic obstructive pulmonary disease, or COPD, is one of the most common lower respiratory chronic diseases. The characteristics of COPD are progressive airflow blockage and breathing-related problems like shortness of breath. Although COPD has mainly negative effects on lung function, significant systemic repercussions have also been recognized.[1] COPD induces system inflammation through increased plasma levels of proinflammatory cytokines and increased circulating inflammatory cells, and can lead to other health complications like coronary artery disease, endothelial dysfunction, depression, altered autonomic nervous system,[2–4] so it has been recognized as multisystem disorder.[5,6]

Although COPD is described as a 'preventable and treatable disease', it is still a global health epidemic with high morbidity and mortality.[7,8,9,10] Data from World Health Organisation put COPD as third leading cause of death worldwide,[11] with an estimated 3.17 million deaths in 2015, which represents 5% of all deaths.[12] The GBD studies estimated 251 million cases of COPD in the world in 2016.[9] Furthermore it is estimated that 4 million people die prematurely from chronic respiratory diseases every year.[13] In The European Union (EU) respiratory diseases are the cause of death in about 600,000 people every year, with more than a half of them from cancer and COPD.[14] According to the data from the Institute for Health Metrics and Evaluation, in 2019 COPD was the sixth leading cause of death in Croatia causing 13.2% of all deaths.[15]

Tobacco exposure, i.e. active or passive smoking, is one of the main risk factors for COPD, but other factors have also been determined (exposure to occupational chemical substances, indoor and outdoor air pollution, genetics, health in early life, previous respiratory infections, nutrition, gender, and socioeconomic status).[14,16–29] A large part of the world's population is exposed to external risk factors for chronic respiratory diseases. Nearly 2.7 billion people in the world rely on biomass, coal or kerosene for cooking,[30] 91% of the world population is exposed to levels of air pollution above WHO Air Quality Guidelines limits,[31,32] and in 2015,

* Corresponding author. Rockefellerova 7, 10 000, Zagreb, Croatia. Tel.: +38514863357.
E-mail address: diana.jovicic@hzjz.hr (D. Jovičić Burić).

https://doi.org/10.1016/j.puhe.2021.10.016
0033-3506/

the age-standardized prevalence of daily smoking worldwide was 25.0% in men and 5.4% in women.[33]

As worldwide, the problem of smoking is present in Croatia, as well as the age-standardized prevalence of daily smoking in 2015 in Croatia was 30.4% in men and 25.9% in women.[33]

Data about prevalence and mortality from COPD have long noted gender differences, with higher rates among men.[12,29] However, it seems that the gender gap is narrowing.[34] Further, epidemiological data show an increase in COPD mortality with age, especially after age 45.[29] Therefore, due to high tobacco smoking rates and population aging in many countries, it can be expected that the burden from COPD would increase even more.[12]

As many external risk factors have an impact on development and mortality from COPD, valuable data come from analysis at the national level. These analyses are recognized as key markers of the situation of the disease that may help in organizing and improving the overburdened healthcare system.[35]

The aim of this study was to analyze the COPD mortality trends in Croatia for the period 2010–2019 and to identify possible changes and differences by age group and gender.

Methods

Mortality data for the period 2010–2019 were obtained from the Croatian Institute of Public Health based on death certificates. In data analysis were included COPD death cases defined as ICD-10 code J44.0 – J44.9.[36] For calculating crude rates, we used the estimated population in the middle of the year from the Croatian Bureau of Statistics for years 2010–2019.[37] Age-standardized mortality rates (ASR) of COPD mortality in Croatia were calculated for each gender, 10-year age group (from 40 years above) and for all ages by the direct standardization method, using the revised EU Standard Population (based on the EU and EFTA 2011–30 population projections) as a reference.[38] These were expressed as rates per 100,000 persons.

To model temporal changes in mortality rates, we carried out joinpoint regression analysis using the Joinpoint Regression Software (Version 4.8.0.1) from the Surveillance Research Program of the US National Cancer Institute.[39] The aim of the approach is to fit the rates of events into the simplest model possible, starting with the model with zero joinpoints. Using the Monte Carlo Permutation method possible, significant joinpoints (breakpoints that would improve the fitness of the model) are identified with a statistical significance set at $P < 0.05$.

For quantifying the trend (for the period between two joinpoints or the beginning and the end of series), a log-linear model with annual percent change (APC, the estimated annual change in rate from one joinpoint to the next in percentage) is calculated with its corresponding 95% confidence intervals (CI).

Statistically significant increase/decrease was considered if the APC was different than zero with $P < 0.05$ and CI not including zero. All statistical tests were two sided.

Results

From 2010 to 2019, there were overall 16,560 COPD death cases in Croatia, 9987 males and 6573 females. The number of deaths ranged from 878 to 1083 annually for men and from 520 to 773 for women. Below age 40, there were only 12 reported COPD death cases in Croatia in the 10-year period.

In both men and women, age group 80+ years old, with 8726 deaths, had the highest number of deaths among the analyzed 10-year age groups (from 40 years above). This age group contributed to 52.7% of all deaths from COPD (45.9% among men and 63% among women) at the national level in the 10-year period analyzed in the study.

Mortality trend analysis – all ages, by gender

Over the 10-year period, there was a stable (not statistically significant decrease) age-standardized COPD mortality rate among men and statistically significant increasing age-standardized COPD mortality rate among women at the national level.

The number of COPD deaths increased in men from 878 in 2010 to 1083 in 2019 and in women from 520 in 2010 to 737 in 2019 (Table 1).

COPD age-standardized mortality trend showed a non-significant decrease for men, with APC -0.24% (95% CI, −1.6% to 1.1%) and a significant increase for women, with APC of 2.93% (95% CI, 0.4% to 5.5%). No joinpoints were identified (Fig. 1).

Mortality trend analysis - 10-year age groups, by gender

In five analyzed 10-year age groups (40–49, 50–59, 60–69, 70–79, 80+) ASR from COPD, for both men and women, there were no joinpoints identified (Fig. 2).

For men age-group 40–49, 50–59, and 70–79, there was a non-significant decreasing age-standardized COPD mortality rate trend. APC for the age group 40–49 was −2.27% (95% CI, −10.2% to 6.4%), for the age group 50–59 was −1.70% (95% CI, −5.2% to 1.9%). A statistically significant decrease by joinpoint analysis was found pointed for men age group 70–79, with APC −1.31% (95% CI, −2.6% to −0.0%); however, as confidence interval includes zero, we consider it non-significant.

In age groups 60–69 and 80+, COPD mortality trend showed a non-significant increase for men, with APC 2.10% for age group 60–69 and 0.33% for age group 80+ (95% CI, −0.5% to 4.7% and −1.5% to 2.2% respectively) (Fig. 2).

For women only age group 50–59 showed decreasing but non-significant COPD mortality trend, with APC -0.43% (95% CI, −4.0% to 3.2%). In age groups 40–49 and 70–79, COPD mortality trend showed a non-significant increase for women, with APC 2.50% for age group 40–49 and 1.27% for age group 70–79 (95% CI, and −1.1% to 3.7% respectively). Significant increasing COPD mortality trend among women at the national level was found for age groups 60–69 and 80+, with APC 5.06% for the age group 60–69 and 3.48% for the age group 80+ (95% CI, 0.9% to 9.4% and 0.7% to 6.3% respectively) (Fig. 2).

Table 1
COPD mortality in Croatian men and women, 2010–2019.

Year	Men			Women		
	N[a]	Crude rate	ASR[b]	N[a]	Crude rate	ASR[b]
2010	878	41.19	68.43	520	22.75	23.12
2011	944	45.73	70.60	549	24.82	23.46
2012	941	45.71	69.97	591	20.46	19.86
2013	977	48.18	79.93	627	28.48	25.73
2014	1019	49.81	73.44	631	28.78	25.66
2015	1083	53.39	74.85	760	34.94	30.37
2016	964	47.85	67.13	684	31.67	27.00
2017	1051	52.81	71.63	773	36.22	29.83
2018	1047	53.01	68.81	701	33.18	26.74
2019	1083	54.96	70.29	737	35.19	27.54

[a] N – number of deaths.
[b] ASR – age-standardized mortality rate per 100,000 (using revised European standard population).

Discussion

This study gave a systematic overview of COPD mortality in Croatia, considering the most important confounding factors - gender and age. Throughout 2010–2019 in Croatia, gender differences in COPD mortality rates were found regardless of age. COPD mortality rates among men remained stable, while among women have increased, especially for ages 60–69 and 80+. On the other hand, a study about COPD mortality rates throughout 1995–2017 showed increased rates among Croatian men, likewise in Czechian men, while rates in Hungary remained stable (since 2005) and in the other European countries, like Italy, Spain, Lithuania, or Finland they were declining.[40] Found narrowing of the gender gap in our study is in line with previous results in which age-specific rates were analyzed. Further, observed convergence between mortality trends among men and women was previously reported in Europe, mostly in high-income countries.[41] As Croatia went through transition and development, becoming EU member in 2013 and a high-income country in 2018, this may be a partial explanation for changes in COPD mortality trend among Croatian men.

The results of this study showed increased COPD mortality in women. This can be due to the increase of the prevalence of COPD in women as the increasing burden of COPD in women has been noted in many countries, i.e. Denmark, England.[42,43]

However, COPD is still considered a disease that primarily affects men. Because that women are at greater risk than men of undiagnosed COPD or misdiagnosed COPD with asthma.[44–47]

As a result, COPD may be well advanced before adequate treatment starts, which increases the risk of complications and lowers the quality of life.

Increased COPD mortality rates among Croatian women could also be explained by an increase in the prevalence of cigarette smoking among women.[48] Given the reduction in gender differences when it comes to this risky behavior among adolescents, an increase in mortality from COPD in women can be expected in the future.[49,50] Some studies found no difference between men and women of the same smoking habits and risk for development of the COPD[51] but some indicate a possible greater susceptibility of women to the harmful effects of smoking on the lungs.[52]

To facilitate early recognition of COPD in women, primary care physicians and other health professionals should be aware of gender differences in the presentation of COPD (e.g. younger age, lower BMI, greater risk for lung impairment, and severe dyspnea for the same level of smoke exposure, and lower socioeconomic status in women in comparison to men).[53–55]

Since our results showed that COPD mortality rates among men in comparison to women remained higher (even more than double), despite the marked convergence, this confirms that men are still at a higher risk of dying from COPD.

Thus, further aims for raising awareness of COPD and health literacy should be focused on both men and women. Health literacy is recognized as an important factor in the prevention and control of non-communicable diseases. It includes access to health services, understanding, appraising, and ability to remember and apply information about health.[56,57] Health literacy has been shown to impact overall health and the use of screening, so patients with lower health literacy present in later stages of the disease, have difficulties in understanding their treatment, which results with lower adherence to medical regimens and higher risk for being hospitalized. Regarding COPD, although it is a disease with high prevalence, studies show there is still poor knowledge of the general population about the disease and its symptoms.[58,59] Therefore, there is a need for interventions aimed at raising public awareness of COPD, and they have been shown to have a positive impact.[60] This positive effect go beyond mentioned use of screening and treatment adherence but also there has been found a positive effect on willingness to quit smoking. Korean study in smoking cessation program showed that after being informed about COPD, their willingness to quit smoking increased.[61]

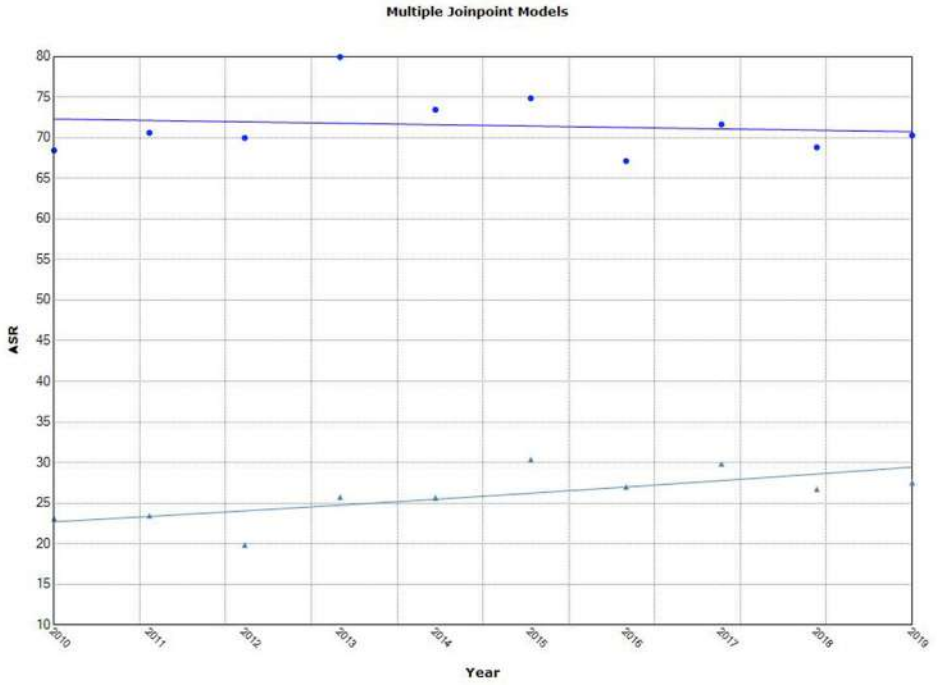


Fig. 1. Joinpoint analysis for COPD mortality in Croatia, by gender, 2010–2019. Circles (upper line) – men; triangles (lower line) - women; ASR – age-standardized mortality rate per 100,000 (using revised European standard population).

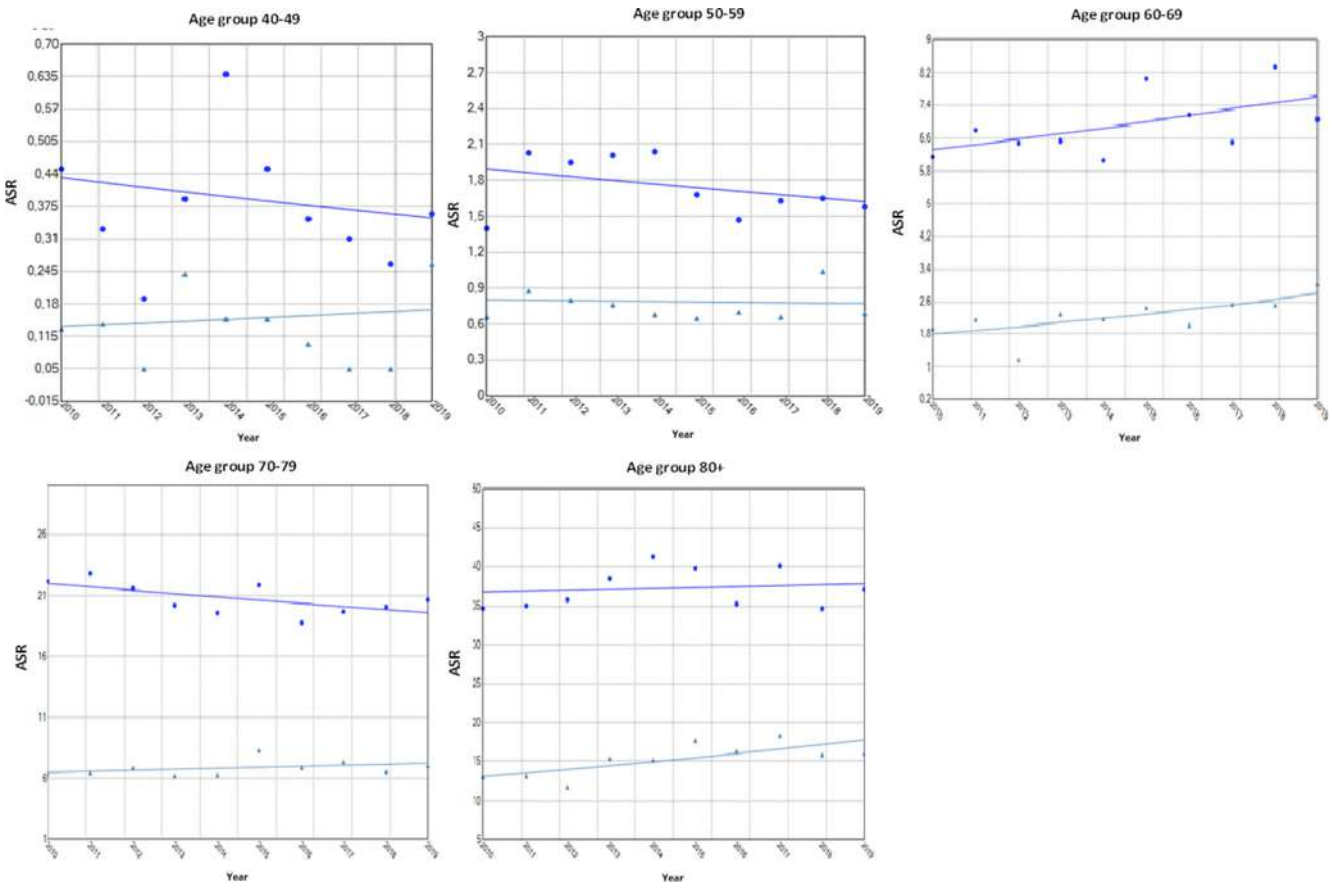


Fig. 2. Joinpoint analysis for COPD mortality in Croatia, by 10-year age groups and gender, 2010–2019. Circles (upper line) – men; triangles (lower line) - women; ASR – age-standardized mortality rate per 100,000 (using revised European standard population).

Further, observed higher COPD mortality rates with age in both men and women confirm previous data and imply that the number of COPD deaths will continue to increase in the future.[12,40] As it was previously found, even with declining COPD mortality rates, the number of COPD deaths has actually been increasing in the majority of the studied countries. Several factors have been recognized. One of the main factors is population ageing.[38] Also, with urbanization, more people are exposed to outdoor air pollution that is a risk factor for COPD development,[51,62] changes in the mortality of other causes such as from cardiovascular disease and acute infection could 'spill over' to COPD mortality.[63]

The low rate of influenza vaccination among the elderly may also be one of the factors as influenza vaccination has been recognized as an important tool for reducing the number of exacerbations, risk of serious illness, and death in patients with COPD.[64] Despite the fact that in Croatia people over 65, as well as those with chronic diseases are provided with free influenza vaccination the prevalence rate of vaccinated among the elderly is low.[65] According to the EUROSTAT in 2018 Croatia had one of the lowest prevalence of people over 65 years old that were vaccinated against influenza.[66] Data from a study conducted in Vukovarsko-Srijemska county in Croatia showed that more than half of the respondents over the age 65 are not vaccinated against influenza regularly every year, and the vast majority of respondents do not want to be vaccinated in the future.[67]

Strengths and limitations

One of the main strengths of this study is that the analysis was conducted based on national COPD mortality data. Mortality data and statistics provide one of the most quality and reliable health data. In Croatia the Tenth Revision of the International Classification of Diseases (ICD-10) rules for determining and coding the underlying cause of death and has been applied since 1995. Since 2012, the Second Edition of the ICD-10 revision has been in force. Also, in line with the recommendations of the WHO and EUROSTAT updates of the ICD-10 revision have been applied since 2004 in order to improve the quality and harmonization of mortality statistics data.[68]

Nevertheless, interpretation of the results should be made while considering several limitations. Because the comorbidity diagnoses such as cancer, cerebrovascular disease, or diabetes are often chosen as the cause of death, although the deceased person also suffered from COPD, recorded COPD mortality may deviate from the actual.[69–73] In addition, unrecognized and undiagnosed COPD in deceased persons or insufficient availability of medical records of the deceased person are also possible reasons for not listing COPD as a cause of death.[72] Therefore, reported COPD mortality rates are likely underestimated. For assessing, more accurately, the real contribution of COPD on mortality, future studies may include contributing causes of death mentioned on the death certificates into the analysis.[40]

Conclusion

This study showed stable COPD mortality rates among men in Croatia and an increase among women, especially for ages 60–69 and 80+. To better understand this change future research should expand analysis and compare national trends with those at the European level.

Mortality indicators are recognized as one of the most valuable for planning and organization of the healthcare system and policies (i.e. assessing the health status of the population, creating health policies, evaluating national health programs, and for regional and international comparisons). Knowing that and with COPD mortality data that indicates that COPD burden will continue to increase in the future, the healthcare system should focus on improvement of the quality of care. Increasing the number of health professionals in prevention, diagnosis, treatment, and rehabilitation would reduce the risk of developing COPD, respectively, reduce the time interval to final diagnosis so that treatment can begin early. This is especially important for the improvement of diagnosing COPD in women. Health professionals should have in mind that women with symptoms of COPD are at higher risk for misdiagnosed COPD with asthma than men with symptoms of COPD. This would reduce mortality and improve the quality of life of patients.

Since COPD is a preventable and treatable disease but still not curable, the primary prevention of COPD remains the most important public health tool. At a population level, further investment in health promotion and prevention programs aimed at reducing risk factors for COPD, especially tobacco smoking, as well as raising awareness and knowledge about COPD as a chronic disease is needed.

Author statements

Ethical approval

For this type of study, formal consent is not required. Data for the analysis was obtained from the CIPH database of death certificates but included only data about age, gender, year of death, and the ICD-10 code of death cause. There was no contact with the individuals.

Funding

None.

Competing interests

None declared.

References

1. Poulain M, Doucet M, Drapeau V, Fournier G, Tremblay A, Poirier P, et al. Metabolic and inflammatory profile in obese patients with chronic obstructive pulmonary disease. *Chron Respir Dis* 2008;**5**:35–41. https://doi.org/10.1177/1479972307087205.
2. Agusti AGN, Noguera A, Sauleda J, Sala E, Pons J, Busquets X. Systemic effects of chronic obstructive pulmonary disease. *Eur Respir J* 2003;**21**:347–60. https://doi.org/10.1183/09031936.03.00405703.
3. Dourado VZ, Tanni SE, Vale SA, Faganello MM, Sanchez FF, Godoy I. Systemic manifestations in chronic obstructive pulmonary disease. *J Bras Pneumol* 2006;**32**. 161-111.
4. Agustí A. Systemic effects of chronic obstructive pulmonary disease: what we know and what we don't know (but should). *Proc Am Thorac Soc* 2007;**4**(7):522–5. https://doi.org/10.1513/pats.200701-004FM.
5. Mannino DM. Chronic obstructive pulmonary disease: definition and epidemiology. *Respir Care* 2003;**48**:1185–91.
6. Agusti À, Soriano JB. COPD as a systemic disease. *COPD: J Chronic Obstructive Pulmonary Dis* 2008;**5**(2):133–8. https://doi.org/10.1080/15412550801941349.
7. Celli BR, MacNee WA, Agusti AA, Anzueto A, Berg B, Buist AS, et al. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respiratory J* 2004;**23**(6):932–46. https://doi.org/10.1183/09031936.04.00014304.
8. Antoljak N. Kronična opstruktivna plućna bolest i astma. *Hrvatski časopis za javno zdravstvo* 2011;**7**(28).
9. Vos T, Abajobir AA, Abate KH, Abbafati C, Abbas KM, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* 2017;**390**(10100):1211–59. https://doi.org/10.1016/S0140-6736(17)32154-2.
10. Barnes PJ. Chronic obstructive pulmonary disease: a growing but neglected global epidemic. *PLoS Med* 2007;**4**(5):e112.
11. WHO Regional Office for Europe [Internet] Available from: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death [cited 2021 Jan 15].
12. [Internet] World Health Organisation. *Chronic obstructive pulmonary disease (COPD)*. Geneva, Switzerland: WHO; 2017. Available from: https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd) [cited 2020 Oct 12].
13. *Forum of International Respiratory Societies the global impact of respiratory disease*. 2nd ed. Sheffield: European Respiratory Society; 2017.
14. [Internet] European Respiratory Society. *Europska plućna bijela knjiga*. 2013. Available from: http://www.erswhitebook.org [cited 2020 Oct 10].
15. Institute for Health Metrics and Evaluation. Croatia [Internet] Available from: http://www.healthdata.org/croatia [cited 2021 Sep 15].
16. Tarjan E, Magyar P, Vaczi Z, Lantos A, Vaszar L. Longitudinal lung function study in heterozygous pimz phenotype subjects. *Eur Respir J* 1994;**7**(12):2199–204. https://doi.org/10.1183/09031936.94.07122199.
17. Lawlor DA, Ebrahim S, Davey Smith G. Association of birth weight with adult lung function: findings from the british women's heart and health study and a meta-analysis. *Thorax* 2005;**60**(10):851–8. https://doi.org/10.1136/thx.2005.042408.
18. Anastassios C, Koumbourlis EKM, Mutich RL, Mallory Jr GB, Walczak SA, Fertal K. Longitudinal follow-up of lung function from childhood to adolescence in prematurely born patients with neonatal chronic lung disease. *Pediatr Pulmonol* 1996;**21**(1):28–34.
19. Li Y-F, Gilliland FD, Berhane K, McConnell R, Gauderman WJ, Rappaport EB, et al. Effects of in utero and environmental tobacco smoke exposure on lung function in boys and girls with and without asthma. *Am J Respir Crit Care Med* 2000;**162**(6):2097–104. https://doi.org/10.1164/ajrccm.162.6.2004178.
20. Strachan D, Cox B, Erzinclioglu S, Walters D, Whichelow M. Ventilatory function and winter fresh fruit consumption in a random sample of british adults. *Thorax* 1991;**46**(9):624–9.
21. Prescott E, Bjerg A, Andersen P, Lange P, Vestbo J. Gender difference in smoking effects on lung function and risk of hospitalization for copd: results from a Danish longitudinal population study. *Eur Respir J* 1997;**10**(4):822–7. https://doi.org/10.1136/thx.46.9.624.
22. Prescott E, Lange P, Vestbo J. Socioeconomic status, lung function and admission to hospital for copd: results from the copenhagen city heart study. *Eur Respir J* 1999;**13**(5):1109–14. https://doi.org/10.1034/j.1399-3003.1999.13e28.x.
23. Adair T, Hoy D, Dettrick Z, Lopez AD. 100 years of mortality due to chronic obstructive pulmonary disease in Australia: the role of tobacco consumption. *Int J Tubercul Lung Dis* 2012;**16**(12):1699–705. https://doi.org/10.5588/ijtld.12.0251.
24. Burney PG, Patel J, Newson R, Minelli C, Naghavi M. Global and regional trends in COPD mortality,1990-2010. *Eur Respir J* 2015;**45**(5):1239–47. https://doi.org/10.1183/09031936.00142414.
25. Burney P, Jithoo A, Kato B, Janson C, Mannino D, Nizankowska-Mogilnicka E, et al. Burden of Obstructive Lung Disease Study. Chronic obstructive pulmonary disease mortality and prevalence: the associations with smoking and poverty–a BOLD analysis. *Thorax* 2014;**69**(5):465–73. https://doi.org/10.1136/thoraxjnl-2013-204460.
26. Taghizadeh N, Vonk JM, Boezen HM. Lifetime smoking history and cause-specific mortality in a cohort study with 43 Years of follow-up. *PLoS One* 2016;**11**(4):e0153310. https://doi.org/10.1371/journal.pone.0153310.
27. Eisner MD, Balmes J, Katz PP, Trupin L, Yelin EH, Blanc PD, et al. Lifetime environmental tobacco smoke exposure and the risk of chronic obstructive pulmonary disease. *Environ Health Perspect* 2005;**4**:7–15. https://doi.org/10.1186/1476-069X-4-7.
28. Sethi JM, Rochester CL. Smoking and chronic obstructive pulmonary disease. *Clin Chest Med* 2000;**21**(1):67–86. https://doi.org/10.1016/s0272-5231(05)70008-3.
29. Raherison C, Girodet PO. Epidemiology of COPD. *Eur Respir Rev* 2009;**18**(114):213–21. https://doi.org/10.1183/09059180.00003609.
30. International Energy Agency. *World energy outlook 2018: highlights*. Paris, France: International Energy Agency; 2018. https://doi.org/10.1787/weo-2018-2-en.
31. World Health Organisation [Internet] Avaliable from: https://www.who.int/health-topics/air-pollution#tab=tab_1 [cited 2021 March 12].
32. Shaddick G, Thomas ML, Mudu P, Ruggeri G, Gumy S. Half the world's population are exposed to increasing air pollution. *NPJ Climate and Atmospheric Science* 2020;**3**(1):1–5.
33. Reitsma MB, Fullman N, Ng M, Salama JS, Abajobir A, Abate KH, et al. Smoking prevalence and attributable disease burden in 195 countries and territories, 1990–2015: a systematic analysis from the Global Burden of Disease Study 2015. *Lancet* 2017;**389**(10082):1885–906.
34. Ntritsos G, Franek J, Belbasis L, Christou MA, Markozannes G, Altman P, et al. Gender-specific estimates of COPD prevalence: a systematic review and meta-analysis. *Int J Chronic Obstr Pulm Dis* 2018;**13**:1507.
35. Lopez-Campos JL, Hartl S, Pozo-Rodriguez F, Roberts CM, European COPD Audit team. Variability of hospital resources for acute care of COPD patients: the European COPD Audit. *Eur Respir J* 2014;**43**:754–62. https://doi.org/10.1183/09031936.00074413.

36. World Health Organization. *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines*. World Health Organization; 1992.
37. Hrvatski Zavod za Statistiku [Croatian Bureau of Statistics] [Internet]. Procjena stanovništva prema dobnim skupinama i spolu u Republici Hrvatskoj. Available from: https://www.dzs.hr [cited 2020 Oct 10].
38. Report of Eurostat's task force *Revision of the European standard population*. 2013 [Internet]. Available from: http://ec.europa.eu/eurostat/documents/3859598/5926869/KS-RA-13-028-EN.PDF/e713fa79-1add-44e8-b23d-5e8fa09b3f8f [cited 2021 May 4].
39. National Cancer Institute. Joinpoint regression program. 3.4.2 ed. Bethesda (MD): Statistical Research and Applications Branch, National Cancer Institute. *2009 Institute NC. Joinpoint trend analysis software (v.4.2.0.1) [EB/OL]*. 2015 [Internet]. Available from: http://surveillance.cancer.gov/joinpoint/ [cited 2015 23 June].
40. Lortet-Tieulent J, Soerjomataram I, López-Campos JL, Ancochea J, Coebergh JW, Soriano JB. International trends in COPD mortality, 1995–2017. *Eur Respir J* 2019;**54**(6). https://doi.org/10.1183/13993003.01791-2019.
41. Lopez-Campos JL, Ruiz-Ramos M, Soriano JB. Mortality trends in chronic obstructive pulmonary disease in Europe, 1994-2010: a joinpoint regression analysis. *Lancet Respir Med* 2014;**2**(1):54–62. https://doi.org/10.1016/S2213-2600(13)70232-7.
42. Bischoff EW, Schermer TR, Bor H, Brown P, van Weel C, van den Bosch WJ. Trends in COPD prevalence and exacerbation rates in Dutch primary care. *Br J Gen Pract* 2009;**59**:927–33.
43. Simpson CR, Hippisley-Cox J, Sheikh A. Trends in the epidemiology of chronic obstructive pulmonary disease in England: a national study of 51804 patients. *Br J Gen Pract* 2010;**60**:277–84.
44. Chapman KR, Tashkin DP, Pye DJ. Gender bias in the diagnosis of COPD. *Chest* 2001;**119**:1691–5.
45. Ancochea J, Miravitlles M, García-Río, Muñoz L, Sánchez G, Sobradillo V, et al. Underdiagnosis of chronic obstructive pulmonary disease in women: quantification of the problem, determinants and proposed actions. *Arch Bronconeumol* 2013;**49**:223–9.
46. Miravitlles M, Soriano JB, García-Río F, Muñoz L, Duran-Tauleria E, Sanchez G, et al. Prevalence of COPD in Spain: impact of undiagnosed COPD on quality of life and daily life activities. *Thorax* 2009;**64**:863–8.
47. Martinez CH, Raparla S, Plauschinat CA, Giardino ND, Rogers B, Beresford J, et al. Gender differences in symptoms and care delivery for chronic obstructive pulmonary disease. *J Womens Health* 2012;**21**:1267–74.
48. Glavak Tkalić R, Miletić GM, Maričić J. *Uporaba sredstava ovisnosti u hrvatskom društvu: istraživanje na općoj populaciji*. Zagreb: Institut društvenih znanosti Ivo Pilar i Ured za suzbijanje zlouporabe droga Vlade Republike Hrvatske. 2016 [Internet]. Available from: https://www.pilar.hr/wp-content/uploads/2018/01/Elaborat_2015.pdf [cited 2021 April 10].
49. Hrvatski zavod za javno zdravstvo i Institut društvenih znanosti Ivo Pilar. *Uporaba sredstava ovisnosti u općoj populaciji Republike Hrvatske: 2019. i analiza trendova uporabe 2011.-2019*. Štimac Grbić, D., Glavak Tkalić, R. (Ur.). Zagreb: Hrvatski zavod za javno zdravstvo i Institut društvenih znanosti Ivo Pilar. 2020 [Internet]. Available from: https://www.hzjz.hr/wp-content/uploads/2021/03/Uporaba_sredstava_ovisnosti_u_opcoj_populaciji_RH_2019.pdf [cited 2021 April 10].
50. ESPAD Group. *ESPAD report 2019: results from the European school survey project on alcohol and other drugs*. Luxembourg: EMCDDA Joint Publications, Publications Office of the European Union; 2020.
51. Mannino DM, Buist AS. Global burden of COPD: risk factors, prevalence, and future trends. *Lancet* 2007;**370**:765–73. https://doi.org/10.1016/S0140-6736(07)61380-4.
52. Sørheim IC, Johannessen A, Gulsvik A, Bakke PS, Silverman EK, DeMeo DL. Gender differences in COPD: are women more susceptible to smoking effects than men? *Thorax* 2010;**65**(6):480–5. https://doi.org/10.1136/thx.2009.122002.
53. Celli B, Vestbo J, Jenkins CR, Jones PW, Ferguson GT, Calverley PM, et al. Sex differences in mortality and clinical expressions of patients with chronic obstructive pulmonary disease. The TORCH experience. *Am J Respir Crit Care Med* 2011;**183**:317–22.
54. Papaioannou AI, Bania E, Alexopoulos EC, Mitsiki E, Malli F, Gourgoulianis KI. Sex discrepancies in COPD patients and burden of the disease in females: a nationwide study in Greece (Greek Obstructive Lung Disease Epidemiology and health ecoNomics: GOLDEN study). *Int J Chronic Obstr Pulm Dis* 2014;**9**:203–13.
55. Jain NK, Thakkar MS, Jain N, Rohan KA, Sharma M. Chronic obstructive pulmonary disease: does gender really matter? *Lung India* 2011;**28**:258–62.
56. World Health Organization (WHO). *WHO independent high-level commission on NCDs report of working group 1*. 2019 [cited 2020 March 10], https://www.who.int/ncds/governance/high-level-commission/HLC2-WG1-report.pdf.
57. Mikkelsen B, Williams J, Rakovac I, Wickramasinghe K, Hennis A, Shin HR, et al. Life course approach to prevention and control of non-communicable diseases. *BMJ* 2019;**28**:364.
58. Soriano JB, Calle M, Montemayor T, Alvarez-Sala JL, Ruiz-Manzano J, Miravitlles M. The general public's knowledge of chronic obstructive pulmonary disease and its determinants: current situation and recent changes. *Arch Bronconeumol* 2012;**48**:308–15.
59. de Queiroz MC, Moreira MA, Jardim JR, Barbosa MA, Minamisava R, Gondim Hdel C, et al. Knowledge about COPD among users of primary health care services. *Int J Chronic Obstr Pulm Dis* 2015;**10**:1–6.
60. Asai M, Tanaka T, Kozu R, Kitagawa C, Tabusadani M, Senjyu H. Effect of a chronic obstructive pulmonary disease (COPD) intervention on COPD awareness in a regional city in Japan. *Intern. Med.* 2015;**54**:163–9.
61. Seo JY, Hwang YI, Mun SY, Kim JH, Kim JH, Park SH, et al. Awareness of COPD in a high risk Korean population. *Yonsei Med J* 2015;**56**:362–7.
62. Lopez-Campos JL, Tan W, Soriano JB. Global burden of COPD. *Respirology* 2016;**21**(1):14–23. https://doi.org/10.1111/resp.12660.
63. Vogelmeier CF, Criner GJ, Martinez FJ, Anzueto A, Barnes PJ, Bourbeau J, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease 2017 report: GOLD executive summary. *Eur Respir J* 2017;**49**(3). https://doi.org/10.1183/13993003.00214-2017.
64. Poole PJ, Chacko E, Wood-Baker RW, Cates CJ. Influenza vaccine for patients with chronic obstructive pulmonary disease. *Cochrane Database Syst Rev* 2006;**1**:CD002733. https://doi.org/10.1002/14651858.CD002733.pub2.
65. Hrvatski zavod za javno zdravstvo [Croatian Institute of Public Health] [Internet]. Available from: https://www.hzjz.hr/sluzba-epidemiologija-zarazne-bolesti/pocinje-cijepljenje-protiv-gripe-za-sezonu-2020-2021/ [cited 2021 April 10].
66. [Internet] Eurostat. *Respiratory diseases statistics*. 2020 [Internet]. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Statistics_Explained:Privacy_policy [cited 2021 March 20].
67. Ropac D, Fridl H, Stašević I. Stavovi starijih osoba prema cijepljenju protiv gripe. *Hrvatski časopis za javno zdravstvo*. 2021:39–50.
68. World Health Organization. *Update and revision committee. Cumulative official updates to ICD-10*. WHO Collaborating Centres for the Family of International Classifications; 2011.
69. Mathers CD, Ma Fat D, Inoue M, Rao C, Lopez AD. Counting the dead and what they died from: an assessment of the global status of cause of death data. *Bull World Health Organ* 2005;**83**:171–7.
70. Ministarstvo zdravstva i socijalne skrbi Republike Hrvatske. Hrvatski zavod za javno zdravstvo. *Priručnik o popunjavanju potvrde o smrti*. 2011. Zagreb.[Internet]. Available from: https://hzjz.hr/wp-content/uploads/2013/11/prirucnik.pdf [cited 2020 Oct 2].
71. Drummond MB, Wise RA, John MT, Zvarich M, McGarvey LP. Accuracy of death certificates in COPD: analysis from the TORCH trial. *COPD: J Chronic Obstructive Pulmonary Dis* 2010;**7**(3):179–85. https://doi.org/10.3109/15412555.2010.481695.
72. Pritt BS, Hardin NJ, Richmond JA, Shapiro SL. Death certification errors at an academic institution. *Arch Pathol Lab Med* 2005;**129**(11):1476–9. https://doi.org/10.1043/1543-2165(2005)129[1476:DCEAAA]2.0.CO;2.
73. Mannino DM, Buist AS, Petty TL, Enright PL, Redd SC. Lung function and mortality in the United States: data from the first national health and nutrition examination survey follow up study. *Thorax* 2003;**58**(5):388–93. https://doi.org/10.1136/thorax.58.5.388.

Public Health 202 (2022) 76–79

Contents lists available at ScienceDirect

Public Health

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

Short Communication

Impact of COVID-19 public health safety measures on births in Scotland between March and May 2020

L.G. Speyer [a, *], L. Marryat [b, c], B. Auyeung [a, d]

[a] School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, United Kingdom
[b] School of Health Sciences, University of Dundee, Dundee, United Kingdom
[c] Salvesen Mindroom Research Centre, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, United Kingdom
[d] Autism Research Centre, Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom

ARTICLE INFO

Article history:
Received 29 January 2021
Received in revised form
1 October 2021
Accepted 29 October 2021
Available online 8 November 2021

Keywords:
COVID-19
Social-distancing
Pregnancy
Birth outcomes

ABSTRACT

Objective: To combat the widespread transmission of COVID-19, many countries, including the United Kingdom, have imposed nationwide lockdowns. Little is known about how these public health safety measures affect pregnant mothers and their offspring. This study aimed to explore the impact of COVID-19 public health safety measures on births in Scotland.
Study design: Cross-sectional study.
Methods: Using routinely collected health data on pregnancy and birth in Scotland, this study compares all births (N = 7342) between 24th March and May 2020 with births in the same period in 2018 (N = 8323) to investigate the potential negative impact of public health safety measures introduced in Scotland in spring 2020. Birth outcomes were compared using Mann-Whitney-U tests and chi-square tests.
Results: Mothers giving birth during the pandemic tended to combine breastfeeding and formula-feeding rather than exclusively breastfeed or exclusively formula-feed, stayed in hospital for fewer days, and more often had an epidural or a spinal anaesthetic compared to women giving birth in 2018.
Conclusion: Overall, results suggest little impact of public health safety measures on birth outcomes. Further research is needed to explore the longer-term impacts of being born in the pandemic on both maternal mental health and child development.

Coronavirus disease 2019 (COVID-19) is spreading rapidly around the globe after its first identification in Wuhan, China, in December of 2019. In response, many countries, including the United Kingdom, have imposed nationwide lockdowns to combat the widespread transmission of COVID-19. In Scotland, the first Covid-19 cases were reported on 1st March 2020, and a strict lockdown was put in place on 24th March. These public health safety measures have had wide ranging effects on everyone, but certain groups, such as pregnant women, might be particularly vulnerable to changes in social contacts and care provisions.[1,2] Pregnant women were ordered to stay at home and self-isolate, partners were only allowed in hospital for the last stages of labour and were not allowed any visitors during their hospital stay. The lessening of parental choice reduced social and formal support, and poorer maternal health compared to prepandemic life may have adverse effects on maternal and neonatal wellbeing. Social-distancing has been shown to lead to an increase in mental health difficulties in the general population[1] and specifically in pregnant women.[2,3] Compared to pre-COVID-19 pregnancy cohorts, women expecting a child during the COVID-19 pandemic in the UK suffer from substantially elevated psychological distress, with 57% reporting clinically relevant symptoms of anxiety, 37% reporting clinically relevant symptoms of depression, and 68% reporting elevated pregnancy-related anxiety.[3] Public health safety measures have further led to a marked rise in domestic violence incidents in the United Kingdom as is reflected in a 49% increase in calls to the national domestic abuse helpline run by the charity Refuge,[4] with pregnant women being of particularly high risk to experience violence also under normal circumstances.[5] As has been shown in prepandemic studies, domestic violence and elevated levels of depression or anxiety in pregnancy are risk factors for adverse maternal and neonatal outcomes.[6]

* Corresponding author. Department of Psychology, University of Cambridge, Downing Site, Cambridge, CB2 3EA, United Kingdom.
E-mail address: ls945@cam.ac.uk (L.G. Speyer).

https://doi.org/10.1016/j.puhe.2021.10.013
0033-3506/

These findings, taken together, highlight the need to investigate the impact of nationwide public health safety measures on pregnancy and birth. This study gives some preliminary evidence on the impact of public health safety measures on 7342 births in Scotland between March and May 2020 using routinely collected health data on pregnancy and birth in Scotland.

The study population comprised 7219 women giving birth to 7342 children (7096 singletons, 246 multiples) in Scotland between 24th March and 31st May 2020, as well as a control group of 8185 women giving birth to 8323 children (8043 singletons, 280 multiples) between March and May 2018. Harmonised routine health data on pregnancy and birth was provided by Public Health Scotland. In particular, obstetric records from the Scottish Morbidity Records (SMR02) were matched with Scottish Birth Records (SBR) and COVID-19 test results from the Electronic Communication of Surveillance in Scotland (ECOSS). Eight women were excluded from analyses as they tested positive for COVID-19 (198 women were tested). While sample sizes for COVID-19 positive women did not allow for further analyses, the data suggested that none of the women or babies had any particularly negative outcome. One important caveat that has to be kept in mind when interpreting the findings of this study is that the cohorts were recruited in two different years. Thus, they may have potentially been exposed to different non-Covid-19 related factors such as changes in health care provisions, which could have influenced the results presented here.

A variety of maternal and infant outcomes were analysed: **induction of labour** (yes, no), **mode of delivery** (unassisted vaginal delivery, planned caesarean section, emergency caesarean section, other (e.g. use of forceps)), **analgesia during labour** (none, gas and air, opioids, epidural, spinal anaesthetic, general anaesthetics, other), **birth outcome** (live birth, stillbirth, infant death), **Appearance, Pulse, Grimace, Activity, and Respiration** (**APGAR**) **score** (low = 0–3, moderately abnormal = 4–6, reassuring = 7–10[7]), **age of gestation, birth weight, length of hospital stay,** and **feeding method on discharge** (breastfeeding, formula, mixed feeding, other). Data were analysed using Mann–Whitney U tests for continuous outcomes and chi-square tests of independence for nominal outcomes. If the chi-square test was significant ($\alpha < 0.05$), posthoc tests (Fisher's exact tests) were conducted to examine all possible comparisons. These were additionally corrected for multiple comparisons using Bonferroni adjustment.

Descriptive Statistics are given in Table 1. Mann–Whitney U tests showed significant results for length of hospital stay with women in 2020 leaving the hospital around 6 hours earlier than women in 2018 ($Z = 9.75$, $P < .001$). There were no significant differences in birth weight ($Z = 0.75$, $P = .454$), or age of gestation ($Z = -0.69$, $P = .488$). Chi-squared tests showed no significant differences in APGAR scores ($\chi^2(2) = 1.28$, $P = .527$), mode of delivery ($\chi^2(3) = 5.53$, $P = .137$), induction of labour ($\chi^2(1) = 0.08$, $P = .783$) and birth outcomes ($\chi^2(2) = 0.60$, $P = .740$), however, there were significant differences in feeding methods on discharge ($\chi^2(3) = 14.70$, $P = .033$) and analgesia during labour and delivery ($\chi^2(6) = 64.56$, $P < .001$). Posthoc tests revealed that women were more likely to combine breastfeeding with formula-feeding (13.2% in 2018 vs 14.8% in 2020) than to exclusively breastfeed (43.4% in 2018 vs 42.9% in 2020, $P = 0.011$, $P^{adj} = .069$) or exclusively formula-feed (42.8% in 2018 vs 41.8% in 2020, $P = 0.006$, $P^{adj} = .038$) Women in 2020 were also more likely to require spinal anaesthetics (29.1% in 2018 vs 33.1% in 2020) compared to using no pain relief air ($P = 0.011$, $P^{adj} = .226$), gas and air ($P = 0.001$, $P^{adj} < .001$) or opioids ($P < 0.001$, $P^{adj} < .001$), as well as more likely to have an epidural (17.6% in 2018 vs 19.9% in 2020) compared to using gas and air ($P < 0.001$, $P^{adj} < .001$) or opioids ($P < 0.001$, $P^{adj} < .001$).

Overall, results suggest that the public health safety measures implemented in response to the COVID-19 pandemic have had relatively little impact on maternal and neonatal outcomes in Scotland. In line with findings from American hospitals,[8] women giving birth in a Scottish hospital during the pandemic tended to leave maternity wards slightly faster than women who gave birth in the same months of 2018. This reduction in hospital stay duration is likely the result of policy modifications that were implemented to protect women, as well as hospital staff, against COVID-19 infections. Birth partners having to leave the hospital right after delivery and limited visitor numbers likely prompted women to go home as soon as possible. There has been some concern that a reduction in hospital stays could lead to increases in the rate of adverse neonatal and maternal outcomes. However, in agreement with other studies looking at the impact of reducing hospital stay durations, our results do not support these concerns.[8]

Women giving birth during the COVID-19 pandemic in Scotland were further found to be more likely to combine breastfeeding with formula-feeding rather than to exclusively breastfeed or exclusively formula-feed. There has been some evidence from British hospitals that women giving birth during the pandemic were more likely to exclusively breastfeed than prepandemic cohorts.[9] This has been attributed to women having more time for themselves and their new-born as they had more help from their partners once home and fewer visitors. There has, however, also been some evidence that women were less likely to continue breastfeeding long term due to a reduction of face-to-face services for breastfeeding support.[10] The increase in mixed feeding that was found in the current study could, however, also be the result of a more general change in feeding practises that is unrelated to the pandemic. For instance, it is possible that more hospitals are now encouraging mothers to supplement breastfeeding with bottle-feeding to counteract infant weight loss, which otherwise puts a lot of pressure on women who may struggle with producing enough breastmilk to exclusively breastfeed. This is, however, purely speculative, and further research is needed to investigate general trends in infant feeding practices.

Results further indicated that women giving birth between March and May 2020 more often had an epidural or received spinal anaesthetics than women giving birth in the same period in 2018. One potential reason for this finding is that birth partners were restricted to just one person who often was only allowed into the labour ward once the expectant mum was already in active labour. This could have resulted in women having reduced pain tolerance in active labour as they were left to cope with the pain of early labour without a supportive birth partner present. Another potential reason for this finding is that during COVID-19, an increased number of consultants and anaesthetic staff were present to provide care for women that may have presented with COVID-19. Thus, this could have made it easier for women to receive an epidural or spinal anaesthetic. However, it is also possible that epidurals and spinal anaesthetics are gaining in popularity independently of the COVID-19 pandemic.

In conclusion, the findings of the current study suggest that public health safety measures implemented in Scotland as a response to the COVID-19 pandemic had a limited impact on maternal and neonatal outcomes. While these findings are reassuring, future research is needed to gain better insights into the impact of COVID-19 and associated public health safety measures on maternal and child health.

Table 1
Descriptive Statistics.

Continuous Variables													
	Cohort 2018				Cohort 2020				Overall				
	M	Median	SD	Range	M	Median	SD	Range	M	Median	SD	Range	*P*
Maternal Age	30.02	30	5.60	14–53	30.34	31	5.49	15–53	30.17	30	5.55	14–53	<.001
Length of Hospital Stay (days)	2.56	2	2.38	0–38	2.29	2	2.50	0–84	2.43	2	.2.44	0–84	<.001
Birth weight	3318.19	3360	608.24	610–5640	3325.73	3374	603.70	620–6000	3321.73	3368	606.11	610–6000	.454
Age of Gestation	38.74	39	2.17	21–44	38.76	39	2.18	22–42	38.75	39	2.17	21–44	.488

Categorical Variables									
	Category	Cohort 2018		Cohort 2020		Overall		Relative Risk Ratio	*P*
		N	%	*N*	%	*N*	%		
Number of Births	Singleton	8043	96.6	7096	96.6	15139	96.7	Reference Group	.998
	Multiples	280	3.3	246	3.4	526	3.3	1.00 (0.99–1.01)	
Sex	Male	4218	50.7	3703	50.4	7921	50.6	Reference Group	.734
	Female	4079	49.3	3638	49.6	7735	49.4	1.01 (0.98–1.04)	
Mode of Delivery	Unassisted Vaginal	4344	52.5	3788	51.6	8132	52.3	Reference Group	.137
	Planned C-Section	1367	16.5	1288	17.5	2655	17.0	1.02 (0.99–1.04)	
	Emergency C-Section	1488	18.0	1363	18.6	2851	18.3	1.01 (0.99–1.04)	
	Other (e.g. Forceps)	1079	13.0	901	12.3	1980	12.7	0.99 (0.97–1.01)	
Induction of Labour	No	5541	67.4	4910	67.2	10451	67.3	Reference Group	.783
	Yes	2678	32.6	2398	32.8	5075	32.7	1.00 (0.98–1.03)	
Birth Outcome	Alive	8273	99.5	7292	99.4	15565	99.4	Reference Group	.740
	Stillbirth	10	0.1	9	0.1	19	0.1	1.00 (1.00–1.00)	
	Infant Death	36	0.4	38	0.5	74	0.5	1.00 (1.00–1.00)	
APGAR Score	0–3	39	0.5	31	0.4	70	0.5	Reference Group	.527
	4–6	144	1.8	112	1.6	256	1.7	0.98 (0.65–1.49)	
	7–10	7959	97.8	7063	98.0	15022	97.9	1.11 (0.69–1.78)	
Method of Feeding at Discharge	Breastfed	3505	43.4	3066	42.9	6571	43.2	Reference Group	.033
	Formula-fed	3452	42.8	2990	41.8	6442	42.3	1.00 (0.96–1.03)	
	Mixed	1066	13.2	1059	14.8	2125	14.0	1.03 (1.01–1.06)	
	Other	54	0.6	33	0.5	78	0.5	1.00 (0.99–1.01)	
Analgesia during Labour and Delivery	None	328	4.7	285	4.3	613	4.5	Reference Group	<.001
	Epidural	1230	17.6	1324	19.9	2554	18.7	1.19 (1.03–1.37)	
	Opioids	1129	16.2	852	12.8	2981	14.5	0.90 (0.78–1.03)	
	Gas and Air	1908	27.4	1710	25.7	3618	26.5	1.03 (0.89–1.19)	
	General Anaesthetics	160	2.3	119	1.8	279	2.0	0.95 (0.87–1.04)	
	Spinal Anaesthetics	2032	29.1	2204	33.1	4236	31.1	1.21 (1.05–1.41)	
	Other	168	2.7	167	2.5	353	2.6	1.05 (0.95–1.15)	

Note. Relative Risk Ratios are given in comparison to the reference group, e.g., breastfed vs formula-fed and breastfed vs mixed with higher/lower ratios indicating that an outcome was more/less likely in the 2020 cohort than in the 2018 cohort, *P*-values are based on Mann–Whitney U tests for continuous variables and on chi-square tests for categorical variables.

Author statements

Acknowledgements

The authors would like to acknowledge the support of the eDRIS Team (Public Health Scotland) for their involvement in obtaining approvals, provisioning and linking data and the use of the secure analytical platform within the National Safe Haven.

Ethical approval

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. Access to the data was granted through the 'electronic Data Research and Innovation Service' (eDRIS) of Public Health Scotland who ensured that our use of the data and analyses would not breach privacy and confidentiality guidelines. The project was approved by the Public Benefit and Privacy Panel for Health and Social Care (HSC-PBPP) and further received approval from the ethics committee at the University of Edinburgh's School of Philosophy, Psychology and Language Sciences (Ref No: 277-1920/1).

Funding

This project was funded by the Data-Driven Innovation (DDI) initiative through the University of Edinburgh as part of their open call for COVID-19 response projects. This research used data assets made available by National Safe Haven as part of the Data and Connectivity National Core Study, led by Health Data Research UK in partnership with the Office for National Statistics and funded by UK Research and Innovation (research which commenced between 1st October 2020 - 31st March 2021 grant ref MC_PC_20029; 1st April 2021 -30th September 2022 grant ref MC_PC_20058).

Competing interests

None declared.

References

1. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China. *Int J Environ Res Public Health* 2020;**17**(5):1729.
2. Ceulemans M, Hompes T, Foulon V. Mental health status of pregnant and breastfeeding women during the COVID-19 pandemic: a call for action. *Int J Gynecol Obstet* 2020 Jul 23;**151**(1) [Internet]. [cited 2020 Nov 12]:

ijgo.13295. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/ijgo.13295.
3. Lebel C, MacKinnon A, Bagshawe M, Tomfohr-Madsen L, Giesbrecht G. *Elevated depression and anxiety among pregnant individuals during the COVID-19 pandemic*. 2020.
4. Domestic abuse and risks of harm within the home [Internet]. Home office preparedness for Covid-19 (coronavirus): domestic abuse and risks of harm within the home - Home Affairs Committee - House of Commons. Publications.parliament.UK.; Available from: https://publications.parliament.uk/pa/cm5801/cmselect/cmhaff/321/32105.htm.
5. Johnson JK, Haider F, Ellis K, Hay DM, Lindow SW. The prevalence of domestic violence in pregnant women. *BJOG An Int J Obstet Gynaecol* 2003;**110**(3):272–5.
6. Schetter CD, Tanner L. Anxiety, depression and stress in pregnancy: implications for mothers, children, research, and practice. *Curr Opin Psychiatry* 2012;**25**(2):141.
7. Simon LV, Hashmi MF, Bragg BN. *APGAR score*. StatPearls [Internet]; 2021 Aug 14 [cited 2021 Sep 28];2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK470569/.
8. Greene NH, Kilpatrick SJ, Wong MS, Ozimek JA, Naqvi M. Impact of labor and delivery unit policy modifications on maternal and neonatal outcomes during the coronavirus disease 2019 pandemic. *Am J Obstet Gynecol MFM* 2020 Nov 1;**2**(4):100234.
9. bhrhospitals.nhs.uk. *Visitor restrictions give more mums the confidence to breastfeed their newborns.* | Latest news |. BHR Hospitals [Internet]; 2020 [cited 2020 Nov 12]. Available from: https://www.bhrhospitals.nhs.uk/news/visitor-restrictions-give-more-mums-the-confidence-to-breastfeed-their-newborns-2625/.
10. Vazquez-Vazquez A, Dib S, Rougeaux E, Wells JC, Fewtrell MS. The impact of the Covid-19 lockdown on the experiences and feeding practices of new mothers in the UK: preliminary data from the COVID-19 New Mum Study. *Appetite* 2020 Jan 1;**156**:104985.

Public Health 202 (2022) 131–138

Contents lists available at ScienceDirect

Public Health

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

Review Paper

Incidence of infectious diseases after earthquakes: a systematic review and meta-analysis

Sara Najafi [a, f], Arash Akahavan Rezayat [a, b, f], Seyyedeh Faezeh Beyzaei [a], Zahra Shahriari [a], Mahdieh Taheri tabar [a], Mohammad Ghasemi Nour [a], Reza Mosaed [c, d], Majid Khadem-Rezaiyan [e], Ramin Hamidi Farahani [b, *]

[a] Student Research Committee, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran
[b] Infectious Disease Research Center, AJA University of Medical Sciences, Tehran, Iran
[c] Department of Clinical Pharmacy, School of Medicine, AJA University of Medical Sciences, Tehran, Iran
[d] Department of Clinical Pharmacy, Faculty of Pharmacy, Tehran University of Medical Sciences, Tehran, Iran
[e] Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

ARTICLE INFO

Article history:
Received 18 December 2020
Received in revised form
13 October 2021
Accepted 8 November 2021
Available online 23 December 2021

Keywords:
Earthquake
Infectious disease
Incidence
Gastrointestinal infection
Dermal infection
Respiratory infection
Central nervous system infection

ABSTRACT

Objectives: Evaluation of the incidence of infectious diseases after natural disasters can help develop healthcare policies. This study provides a global review of the most prevalent infectious diseases observed after earthquakes.
Study design: A systematic review and meta-analysis were performed.
Methods: A systematic review was performed on electronic databases, including PubMed, Scopus and Web of Science, up to March 2020 (with no time limitations). Studies addressing earthquakes and infectious diseases were collected based on specified inclusion and exclusion criteria. Subsequently, the quality of the studies was assessed by the Newcastle-Ottawa scale (NOS). Data analyses were carried out on six subgroups under five different disease categories using comprehensive meta-analysis software.
Results: In total, 24 studies qualified for the systematic review and 18 were included in the meta-analysis. The incidences of gastrointestinal infections, dermal infections, respiratory infections, central nervous system infections and other infectious diseases were as follows: odds ratio (OR) 163.4 (95% confidence interval [CI]: 31.0–858.1), OR 84.5 (95% CI: 27.1–262.8), OR 9.9 (95% CI: 3.5–27.7), OR 0.5 (95% CI: 0.2–1.1) and OR 4.4 (95% CI: 1.9–9.9) cases per 100,000 people, respectively. The association between the incidences of infectious diseases before and after earthquakes was significant, namely, 1.561 (95% CI: 1.244–1.957) with a *P*-value <0.001.
Conclusions: The results show an increase in the prevalence of infectious diseases after earthquakes. Governments should take essential measures to be better prepared for such unpredictable catastrophes.

Introduction

Earthquakes occur throughout the world with considerable annual frequency. However, only a small number of earthquakes are sufficiently strong to be felt and cause damage. Earthquakes occur at focal depths of 700 km under the earth's crust and show a sharp decrease in destructive strength with increasing distance from the epicentre.[1] Earthquakes are among the most frequent natural disasters and were responsible for approximately 1.87 million deaths in the 20th century.[2] Based on the National Comorbidity Survey, 18.9% of men and 15.2% of women reported a lifetime experience of a natural disaster.[3] Natural disasters may result in the outbreak of infectious diseases as a result of extensive population displacement. Natural disasters can also increase the risk factors for disease transmission, such as variation in the environment, human conditions and susceptibility to pathogens.[4] People who live in affected areas are usually forced to change their lifestyles because of scarce water availability and deficiency of food supplies. In addition, natural disasters impact medical treatment through the loss of medicines, damage to hospitals and healthcare facilities, and reduced emergency capacity. Such situations worsen the physical and mental health conditions of injured individuals.[5–8] Standard medical resources in

* Corresponding author. Aja University of Medical Science, Shahid Etemad-Zadeh St, West Fatemi St, Tehran, 1411718541, Iran. Tel.: +98 21 88028350.
E-mail address: r.hamidi@ajaums.ac.ir (R. Hamidi Farahani).
[f] Sara Najafi and Arash Akhavan Rezayat contributed equally to this work.

https://doi.org/10.1016/j.puhe.2021.11.005
0033-3506/

impacted regions may not be able to respond effectively to the crisis; therefore, leading organisations must provide the necessary medical equipment and supplies, which requires planning, in addition to a profound understanding of diseases and risk of infections.[9–11]

To identify post-earthquake infectious disease outbreaks and risk factors of earthquake-related injuries, disabilities and traumas in different populations, a comprehensive search was carried out on the related literature. Such studies are necessary to examine the probable risk factors, such as climate changes, geographic regions and other circumstances that exacerbate the incidence of different infectious diseases. However, no systematic reviews and meta-analyses that specifically addressed post-earthquake infectious diseases were found. This systematic review and meta-analysis provide a clear global insight into the incidences of earthquake-related infectious diseases. In addition, this study classifies diseases and suggests potential areas where more consideration and increased public attention are needed to reduce the risk of these diseases and their corresponding complications.

Methods

Literature search

A systematic review of peer-reviewed, English-language articles was performed by searching indexed electronic databases, including PubMed, Scopus and Web of Science. Two reviewers separately searched the articles published up to March 2020 (with no time limitations), providing information on earthquakes, infectious diseases and their incidences. Google and Google Scholar were also used as sources of grey literature. Moreover, the references of all included articles were manually searched.

The search algorithms were developed based on the syntax of each database regarding infectious diseases, epidemiology, incidence and earthquakes.

Inclusion and exclusion criteria

Observational studies on the incidence of infectious diseases before and after earthquakes were included. Articles about irrelevant types of disasters, non-English articles, reviews, books, theses, letters to editors, unstructured papers, proceeding papers and dissertations were excluded. Articles that did not provide information on the incidence of infectious diseases and the total population and articles with unavailable full-text were also omitted from the review process.

Study selection and screening

Three authors (FB, MT and ZS) searched the databases based on the given search strategy. All original studies were retrieved, combined and exported into EndNote X7 software (Thomson Reuters, New York, NY, USA). Duplicated articles were removed. According to the inclusion and exclusion criteria, three investigators (SN, FB and ZS) independently screened the title, abstract and text of all articles. Two reviewers (FB and SN) evaluated any inconsistencies in the final selected articles. Finally, 24 articles were selected for the study (Fig. 1).

Quality assessment and data extraction

Two independent researchers used the Newcastle-Ottawa Scale (NOS)[12] to evaluate the quality of selected studies in terms of selection, comparability and outcome.[13] The NOS scores ranged from 0 to 10, with 0–3 indicating poor quality, 4–6 indicating intermediate quality and >7 signifying high quality.[13]

The data extraction form was designed based on similar review studies found in the initial search. Three researchers gathered and summarised the information as shown in Table 1. The extracted parameters were as follows: (i) publication details (last name of the first author, publication year and study design); (ii) general characteristics of the earthquake (country, the epicentre and earthquake year); and (iii) details of the infectious diseases (age range of those infected, infectious diseases and incidence of diseases before and after the earthquake). Infectious diseases were classified into four categories: respiratory, gastrointestinal and hepatic, CNS and dermal.

Data analyses

Data from each study were analysed under seven different subgroups using comprehensive meta-analysis software, version 3.3.070 (Biostat, Englewood, NJ, USA). Heterogeneity was measured by the inconsistency index (I2). A random-effect model was performed for I2 >75%. The final data were summarised as frequencies, odds ratios (ORs) and 95% confidence intervals (CIs). The publication bias was assessed by Egger's test and the Trim and Fill method.

Results

Search and selection process

According to the literature search strategies, 1461, 434 and 380 articles were found in Scopus, PubMed and Web of Science, respectively. Moreover, 220 publications were added by searching Google Scholar and reviewing the reference lists of the relevant studies. After removing duplicates and investigating inclusion and exclusion criteria, 1667 publications were excluded. Full texts of the remaining studies ($n = 257$) were assessed. Subsequently, 24 studies were included in the systematic review and 18 were eligible for meta-analysis (Fig. 1 , PRISMA flow chart).

Study characteristics

The majority of studies were cross-sectional ($n = 16$), followed by retrospective studies ($n = 3$), descriptive follow-ups ($n = 2$), epidemiological investigations ($n = 2$) and a one-time series investigation ($n = 1$). Studies were published between 1982 and 2017. All publications scored >5 in the quality assessment. Studies related to the Great East Japan earthquake were the most frequent ($n = 6$). Respiratory infections were investigated most frequently ($n = 13$), and acute respiratory infection was the most common disease to be reviewed. Some studies compared the incidences of pre- and post-earthquake infectious diseases and patient numbers ($n = 7$), whereas other studies only assessed the post-earthquake situation. The total aggregated sample size of eligible studies for meta-analysis was 18 (Table 1 and Fig. 1). Further information about demographic data is available in Supplementary Table1.

Meta-analysis

In total, 18 studies were included in the six-subgroup analysis (Table 2).

Disease categories

Five different disease categories were analysed. The highest incidence was for gastrointestinal and hepatic infection, with 163.4 cases per 100,000 people (95% CI: 31.0–858.1). Supplementary Figure 1 The second highest was for dermal infection, with 84.5 cases per 100,000 people (95% CI: 27.1–262.8), Supplementary Figure 2 followed by respiratory infection with 9.9 cases per 100,000 people (95% CI: 3.5–27.7). Supplementary Figure 3 The fourth was CNS infection, with

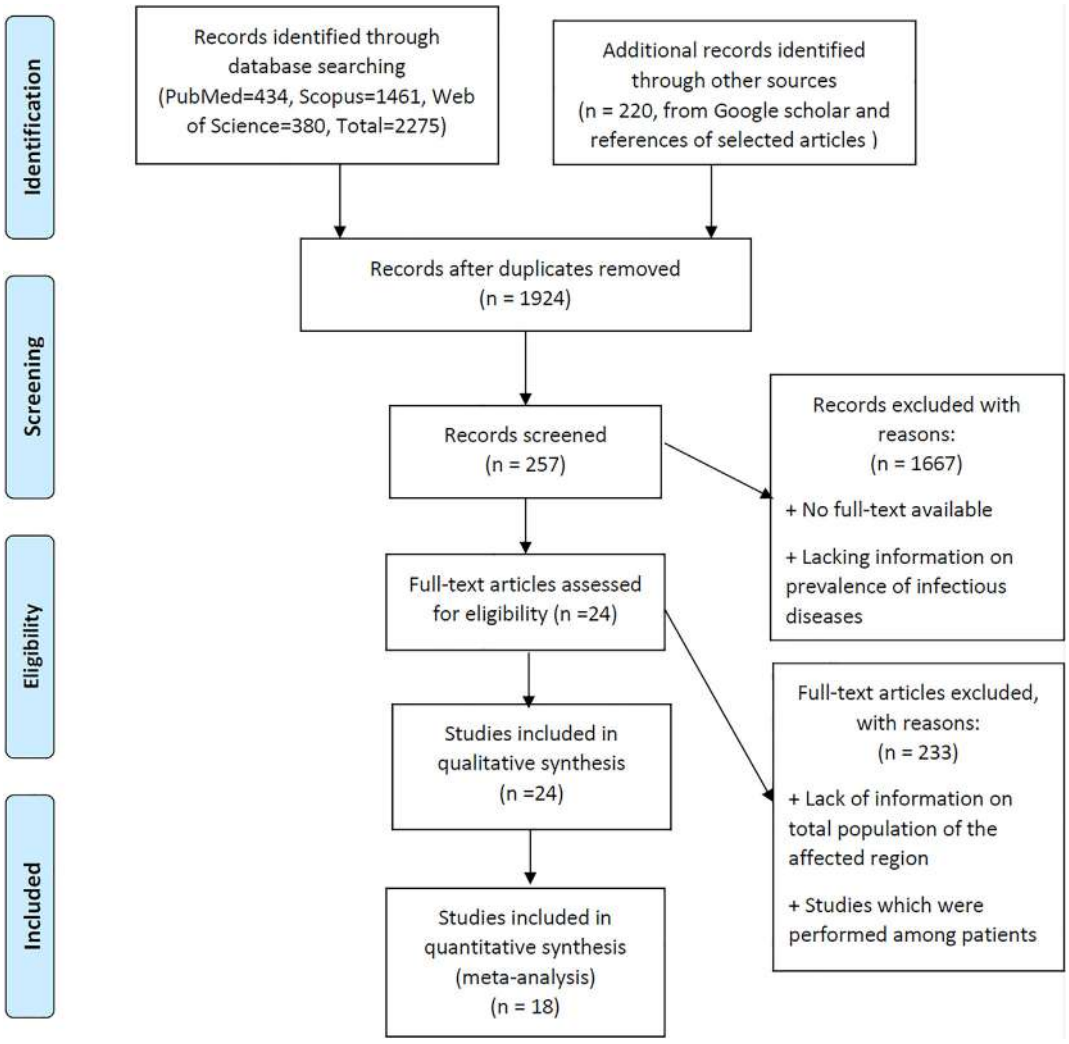


Fig. 1. PRISMA flow diagram.

0.5 cases per 100,000 people (95% CI: 0.2–1.1) Supplementary Figure 4 and 'other diseases' had an incidence rate of 4.4 (95% CI: 1.9–9.9) cases per 100,000 people. Supplementary Figure 5 The assessments of gastrointestinal and hepatic infection, dermal infection, respiratory infection, CNS infection and 'other diseases' were included in 8, 9, 22, 7 and 20 studies, respectively.

The Egger's test and Trim and Fill methods indicated publication bias for respiratory infections (Egger test, $P < 0.001$) and demonstrated no significant publication bias for other disease groups (Supplementary Table2).

Gastrointestinal and hepatic infections:

This category was divided into two leading diseases: viral hepatitis that occurred with an incidence of 456.6 cases per 100,000 people (95% CI: 118.5–1743.4) and diarrhoea with 56.8 cases per 100,000 people (95% CI: 5.6–572.3). Results on the incidence of post-earthquake gastrointestinal and hepatic infections were reported from Asia and Europe.

Dermal infections

Cutaneous leishmaniasis was the most frequently reported dermal infection, with 471.7 cases per 100,000 people (95% CI: 142.8–1546.6). One-third of the reported cases in Asia occurred in the Iran earthquake (2003).

Respiratory infections

Acute respiratory infection was the most common disease in this category, with an incidence of 328.5 cases per 100,000 people (95% CI: 133.3–807.2). Valley fever (in the US), pneumonia, pertussis and tuberculosis scored the subsequent ranks, respectively. In terms of the earthquake time and continent analysis, most results were reported in America, Asia and Europe.

CNS infections

Meningococcal meningitis was the most common CNS infection, with an incidence of 0.4 cases per 100,000 people (95% CI: 0.1–1.4).

Other infections

Typhoid fever, malaria, HIV and German measles are uncategorised diseases that were abundant in America, Europe and Asia.

Pre-earthquake and post-earthquake

Data from seven articles, considering 14 different infectious diseases, were evaluated both before and after earthquakes. The association between the incidences of infectious diseases before and after the earthquakes was significant with an OR of 1.561 (95% CI: 1.244–1.957) and a *P*-value of <0.001 (Table 3 and Fig. 2). Furthermore, dermal infections showed significant differences

Table 1
Characteristic data of the included studies.

First Author	City or county or province, country	Disease category	Disease	Incidence of disease from 1 year before the earthquake (cases/ 100,000)	Incidence of disease until 1 year after the earthquake (cases/ 100,000)
Aflatoonian	Nezamshahr County, Iran	DI	Anthroponotic cutaneous leishmaniasis		1208.51
Alexander[35]	Naples, Italy	RI	Diphtheria		0.08[(2)]
		RI	Pertussis		0.2[(2)]
		GI & HI	Viral hepatitis		56.9[(2)]
		CNSI	Meningococcal meningitis		0.1[(2)]
		OI	Typhoid fever		14.9[(2)]
		OI	German measles		0.2[(2)]
		OI	Paratyphoid		9.15[(2)]
	Region of Basilicata, Italy	RI	Pertussis		3.2[(2)]
		GI & HI	Viral hepatitis		45.9[(2)]
		CNSI	Meningococcal meningitis		1.9[(2)]
		OI	Typhoid fever		16.2[(2)]
		OI	German measles		8.4[(2)]
	Region of Campania, Italy	RI	Pertussis		0.7[(2)]
		GI & HI	Viral hepatitis		53.5[(2)]
		CNSI	Meningococcal meningitis		1[(2)]
		OI	Typhoid fever		18.2[(2)]
		OI	German measles		0.3[(2)]
Ayoagi[14]	Ishinomaki, Kesennuma and Sendai, Japan	RI	Tuberculosis	0.14[(3)]	0[(4)]
		RI	Acute respiratory infections	0.28[(3)]	0.55[(4)]
		RI	Influenza and pneumonia	0.47[(3)]	3.68[(4)]
		GI & HI	Intestinal infectious disease	0.2[(3)]	0.28[(4)]
		CNSI	Tetanus	0[(3)]	0.11[(4)]
		CNSI	Viral infections of the CNS	0.07[(3)]	0[(4)]
		DI	Erysipelas	0.07[(3)]	0[(4)]
		DI	Viral infections characterised by skin and mucous membrane lesions	0[(3)]	0.17[(4)]
		OI	Human Immunodeficiency Virus	0.07[(3)]	0[(4)]
		OI	Infectious arthropathies	0.07[(3)]	0.06[(4)]
		OI	Urinary tract infection	0.14[(3)]	0.33[(4)]
Daito	Kesennuma City, Japan	RI	Pneumonia	481.33[(5)]	340.40[(6)]
Fakoorziba	Zarin dasht, Iran	DI	Cutaneous leishmaniasis	491	864
	Darab, Iran	DI	Cutaneous leishmaniasis	225	218
	Fars, Iran	DI	Cutaneous leishmaniasis	105	117
Feng	Ludian, China	OI	Malaria		0[(7)]
	Yongshan, China	OI	Malaria		0[(7)]
	Jinggu, China	OI	Malaria	0.013[(8)]	0.003[(7)]
Furusawa[36]	Titiana, Solomon Islands	OI	Malaria		546.44[(9)]
	Tapurai, Solomon Islands	OI	Malaria		854.7[(9)]
	Mondo, Solomon Islands	OI	Malaria		586.51[(9)]
	Olive, Solomon Islands	OI	Malaria		821.91[(9)]
Greco[37]	Southern Italy, the epicentre was located in Southern Campania, Italy	RI	Pertussis		0.53[(10)]
		GI & HI	Diarrhoea with fever		5.14[(10)]
		CNSI	Meningococcal meningitis		0.77[(10)]
		OI	Typhoid fever		1.7[(10)]
Jonaidi Jafari[38]	Bam, Iran	RI	Respiratory infection		7120[(11)]
		RI	Tuberculosis		12[(11)]
		GI & HI	Diarrhoea		820[(11)]
		CNSI	Acute flaccid paralysis		0[(11)]
		DI	Anthrax		0[(11)]
		DI	Cutaneous leishmaniasis		75[(11)]
		OI	Acute icterus		3[(11)]
Kamigaki	Yamamoto, Japan	RI	Acute respiratory infection		5801.1[(12)]
Karmakar[39]	Kashmir, Pakistan	RI	Acute respiratory infection		3295.38[(13)]
		GI & HI	Acute diarrhoeal disease		2743.07[(13)]
Kawano	Oukaidou, Oukuma, Japan	RI	Acute respiratory infection		16880[(14)]
		GI & HI	Acute gastroenteritis		2370[(14)]
Kawano	Ishinomaki, Japan	RI	Acute respiratory infection		54[(15)],35[(16)]
Khan[40]	North of Pakistan, Pakistan	GI & HI	Hepatitis C Virus		3265.3[(17)], 5517.24[(18)]
		OI	Human Immunodeficiency Virus		0[(17)]
Matsuoka[41]	Kobe, Japan	RI	Pneumonia		28.4[(19)]
	Ashiya, Japan	RI	Pneumonia		9.2[(19)]
	Nishinomiya, Japan	RI	Pneumonia		15.3[(19)]
	Amagasaki, Japan	RI	Pneumonia		2.8[(19)]
	Itami, Japan	RI	Pneumonia		1.1[(19)]

Table 1 (*continued*)

First Author	City or county or province, country	Disease category	Disease	Incidence of disease from 1 year before the earthquake (cases/100,000)	Incidence of disease until 1 year after the earthquake (cases/100,000)
	Takarazuka, Japan	RI	Pneumonia		1.5[19]
	Kawanishi, Japan	RI	Pneumonia		8.4[19]
	Awaji, Japan	RI	Pneumonia		0.6[19]
Ortiz[42]	Coast of Ecuador, Ecuador	OI	Zika virus	3.57[20]	56.92[21]
Salazar[43]	Bohol, Philippines	OI	Infectious disease		696[23]
Sharifi	Bam, Iran	DI	Anthroponotic cutaneous leishmaniasis		5252
Schneider[44]	Ventura County, USA	RI	Valley fever		30[24]
	Simi Valley, USA	RI	Valley fever		114[24]
Tohma[45]	Tohoku, Japan	RI	Influenza		0.089[25]
Townes	Léogâne and Jacmel, Haiti	OI	Malaria		116.11[26]
Vasquez	Manabí, Ecuador	OI	Zika virus	0.94[27]	10.15[28]
Zhang	Longnan City, China	CNSI	Japanese encephalitis	1.25	1.36
		DI	Kala-azar	4.04	4.4

RI: respiratory infections, GI & HI: gastrointestinal and hepatic infections, DI: dermal infections, CNSI: central nervous system infections and OI: other infections.
1- Incidence of the disease from1978 to1980.
2- Incidence of the disease 3 months after the earthquake.
3- Incidence of the disease in the same month before the earthquake.
4- Incidence of the disease one month after the earthquake.
5- Incidence of the disease among adults over 18 years.
6- Incidence of the disease among adults over 18 years, in 109 days after the disaster.
7- Incidence of the disease 9 months after the earthquake.
8- Incidence of the disease 9 months before the earthquake.
9- Incidence of the disease 2 years after the earthquake.
10- Incidence of the disease about 6 months after the earthquake.
11- Incidence rate of the disease one month after the earthquake.
12- Incidence of the disease 20 days after the earthquake.
13- Incidence of the disease after about 2 months.
14- Incidence of the disease among evacuees in shelters per week in 21 days after the earthquake.
15- Cases per day in crowded shelters until 21 days after the earthquake.
16- Cases per day in a non-crowded shelter until 21 days after the earthquake.
17- Incidence of the disease 2 months after the earthquake.
18- Incidence of the disease 11 months after the earthquake.
19- Incidence of the disease 15 days after the earthquake.
20- Incidence of the disease 18 weeks before the earthquake.
21- Incidence of the disease 18 weeks after the earthquake.
22- Cases per epi-week.
23- Incidence of the disease 150 days after the earthquake.
24- Incidence of the disease 51 days after the earthquake.
25- Incidence of the disease about 2 months after the earthquake.
26- Incidence of the disease about 1 month after the earthquake.
27- Incidence of the disease 104 days before the earthquake.
28- Incidence of the disease 98 days after the earthquake.
29- Incidence of the disease from1978 to1980.
30- Incidence of the disease 3 months after the earthquake.
31- Incidence of the disease in the same month before the earthquake.
32- Incidence of the disease one month after the earthquake.
33- Incidence of the disease among adults over 18 years.
34- Incidence of the disease among adults over 18 years, in 109 days after the disaster.
35- Incidence of the disease 9 months after the earthquake.
36- Incidence of the disease 9 months before the earthquake.
37- Incidence of the disease 2 years after the earthquake.
38- Incidence of the disease about 6 months after the earthquake.
39- Incidence rate of the disease one month after the earthquake.
40- Incidence of the disease 20 days after the earthquake.
41- Incidence of the disease after about 2 months.
42- Incidence of the disease among evacuees in shelters per week in 21 days after the earthquake.
43- Cases per day in crowded shelters until 21 days after the earthquake.
44- Cases per day in a non-crowded shelter until 21 days after the earthquake.
45- Incidence of the disease 2 months after the earthquake.
46- Incidence of the disease 11 months after the earthquake.
47- Incidence of the disease 15 days after the earthquake.
48- Incidence of the disease 18 weeks before the earthquake.
49- Incidence of the disease 18 weeks after the earthquake.
50- Cases per epi-week.
51- Incidence of the disease 150 days after the earthquake.
52- Incidence of the disease 51 days after the earthquake.
53- Incidence of the disease about 2 months after the earthquake.
54- Incidence of the disease about 1 month after the earthquake.
55- Incidence of the disease 104 days before the earthquake.
56- Incidence of the disease 98 days after the earthquake.

Table 2
Meta-analysis of the incidence of infectious diseases in 18 studies. The analysis was conducted based on two subgroups[1]: prevalent diseases in each disease category[2] and distribution in continents for each disease category.

Disease Category	Subgroup	Number of Studies	Event Rate (Cases/100,000)	Lower Limit (Cases/100,000)	Upper Limit (Cases/100,000)	*P*-value
Prevalent diseases						
Respiratory	Acute respiratory infection	3	328.5	133.3	807.2	0.000
	Tuberculosis	2	0.7	0.0	258.6	0.000
	Pneumonia	10	7.0	2.1	22.8	0.000
	Pertussis	3	0.7	0.2	3.1	0.000
	Valley fever	2	58.7	16.1	214.1	0.000
	Others	2	23.1	0.0	93476.2	0.137
	Total	22	9.9	3.5	27.7	0.000
GIT	Viral hepatitis	4	456.6	118.5	1743.4	<0.001
	Diarrhoea & GIT infections	4	56.8	5.6	572.3	0.000
	Total	8	163.4	31.0	858.1	0.000
CNS	Meningococcal meningitis	4	0.4	0.1	1.4	0.000
	Others	3	0.5	0.1	2.9	0.000
	Total	7	0.5	0.2	1.1	0.000
Dermal	Cutaneous leishmaniasis	6	471.7	142.8	1546.6	0.000
	Others	3	0.5	0.0	1.1	<0.001
	Total	9	84.5	27.1	262.8	0.000
Others	Malaria	8	6.2	0.3	131.0	<0.001
	German measles	2	1.4	0.0	53.0	0.000
	Typhoid fever	3	7.4	1.9	29.9	0.000
	HIV	2	1.6	0.0	4393.4	0.007
	Others	5	2.2	0.9	5.2	0.000
	Total	20	4.4	1.9	9.9	0.000
Continent						
Respiratory	Asia	16	17.5	6.0	51.3	0.000
	Europe	4	0.5	0.1	1.9	0.000
	America	2	58.7	16.1	214.1	0.000
GIT	Asia	5	519.5	171.1	1566.4	0.000
	Europe	3	24.2	5.1	113.7	0.000
CNS	Asia	4	0.3	0.0	1.8	0.000
	Europe	3	0.7	0.2	2.1	0.000
Dermal	Asia	9	84.5	27.1	262.8	0.000
Others	Asia	8	0.1	0.0	0.8	0.000
	Europe	10	26.7	12.3	57.9	0.000
	America	2	34.4	3.2	373.6	0.000

CNS, central nervous system and GIT, gastrointestinal.

Table 3
Pre- and post-earthquake disease incidence analysis.

Subgroup	Number of Studies	Odds Ratio	Lower Limit	Upper Limit	Z-value	*P*-value
Respiratory	4	1.321	0.346	5.046	0.407	0.684
GIT	1	1.235	0.698	2.184	0.725	0.468
Dermal	5	1.272	1.007	1.607	2.014	0.044
CNS	2	1.598	0.404	6.325	0.668	0.504
Others	2	4.110	0.625	27.036	1.471	0.141
Total	14	1.561	1.244	1.957	3.852	0.000

CNS, central nervous system and GIT, gastrointestinal.

following earthquakes with an OR of 1.272 (95% CI: 1.007–1.607) and a *P*-value of 0.044.[14–18]

Discussion

This study aimed to investigate the prevalence of infectious diseases after earthquakes around the world. The results show that the overall incidence of infectious diseases increased following an earthquake. The results also indicate that gastrointestinal, dermal and respiratory infections are the most common post-earthquake infectious diseases. The timing and power of an earthquake, its region, the presence of nearby seas and the occurrence of other subsequent natural disasters can play essential roles in the outbreak of infectious diseases.[19,20]

As illustrated, gastrointestinal and hepatic infections were the most prevalent disease category. This is typically due to the inaccessibility of safe drinking water, poor hygiene practices and overcrowding living conditions due to displacement and populated sheltering.[21]

The majority of dermal infections were leishmaniasis, which is an important vector-borne disease.[22] This observation may be explained by the fact that earthquakes cause damages to buildings, and residents may have to stay outdoors, increasing their susceptibility to insect bites.[16,18] Earthquakes can also change animal and vector habitats.[18] Cutaneous leishmaniasis can impact all age groups, but it has been mostly observed in children, probably due to their immature acquired immunity.[16,22,23]

Respiratory infections were another prevalent disease category. Possible explanations for increased respiratory infections could be the lack of appropriate water and personal hygiene, malnutrition, insufficient health services and crowded shelters leading to close contact between individuals.[9,24–26] Freezing temperatures and lack of standard heating equipment could be a prominent cause of post-earthquake pneumonia outbreaks.[15]

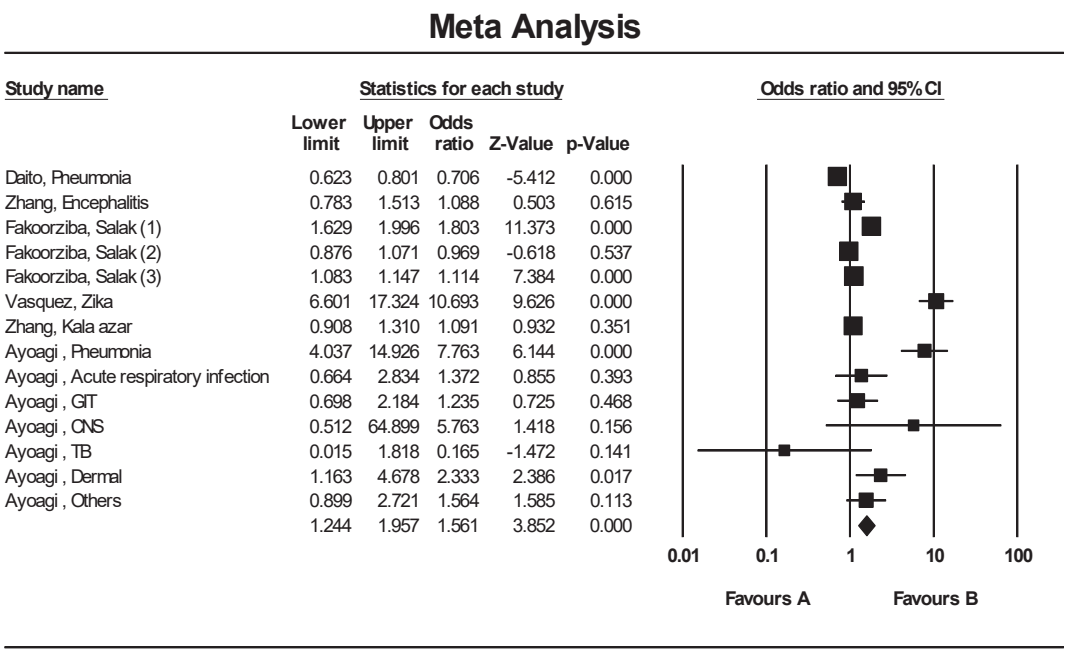


Fig. 2. The Forest plot of comparison of pre-earthquake infections incidence (Favours A) vs post-earthquake infections incidence (Favours B); the last line demonstrates the overall incidence and 95% confidence interval (CI). CNS, central nervous system and GIT, gastrointestinal.

In the current study, within the CNS infectious disease category, the prevalences of Japanese encephalitis and viral infections of the CNS were not significantly different before and after earthquakes. Tetanus could be a potential health threat following injury, as observed after the Great East Japan earthquake.[14]

Among uncategorised infections, malaria and the Zika virus were the most reported diseases. Malaria and Zika virus are vector-borne diseases. Earthquakes can change the habitats of animals and vectors.[27] Living outdoors,[28] crowding of infected and susceptible hosts, damaged health infrastructure, ceasing health programmes and weakened immune systems due to the stressful conditions are among the risk factors of vector-borne disease transmission.[17,27]

According to disease prevention programmes, reinforcement of health surveillance systems, practising guidelines for managing information on specific diseases and increasing the awareness of high-risk populations about communicable diseases and the prerequisites for quick referral to a health facility are highly essential.[29,30] Emergency resource forecasting is one of the mechanisms to increase disaster preparedness. In addition, in countries with the potential threat of disasters, providing fully operational field hospitals, constructing durable buildings and providing effective and efficient healthcare services for vulnerable individuals in the probable forthcoming disasters are critical.[31,32] Food safety is also crucial for disease prevention in natural disasters.

Vector control interventions, based on the local context and epidemiology of diseases, are essential for the prevention of vector-borne diseases. For example, indoor residual spraying for malaria, insecticide-treated nets and traps for tsetse flies could be highly beneficial.[33] The public sector, academics and private organisations should continuously cooperate to resource and fund programmes to provide up-to-date education and training.[34]

Similar to other studies, the current review has some limitations. First, the information from most evaluated articles in this field was outdated, or the patient's medical history was incomplete or was not available. In addition, the number of published articles was inadequate compared to the number of earthquakes and their corresponding problems. Another noticeable limitation was that only a limited number of investigations compared pre-earthquake and post-earthquake incidences of infectious diseases. Therefore, the increase in infectious diseases cannot be precisely assigned to the disaster. On the other hand, the total population of patients and people involved in natural disasters should also be considered. Most articles narrowed their population to a specific community, such as patients referred to a hospital or living within a particular region. In addition, each study only considered a few diseases in their investigations, and the data of some common infectious diseases were not available. Thus, the results of this study are limited to the aforementioned infections. As natural disasters occurred in different areas under different climatic conditions, the results cannot be applied to the global general population, resulting in difficulties in interpretation. Further investigations are required to determine the effects of season and weather on the onset of infectious diseases following earthquakes.

In conclusion, the results of this study show an increase in the incidence of infectious diseases after earthquakes. Gastrointestinal, dermal and respiratory infections were among the most frequently reported diseases. It is recommended that governments have appropriate plans for preventing infectious disease outbreaks both before and after natural disasters. These strategies may include providing vaccination and sufficient emergency healthcare services before the catastrophes, and adequate shelters, clean water, sufficient food supplies and reliable healthcare systems immediately after natural disasters.

Acknowledgment

The authors would like to respectfully thank the Student Research Committee of Mashhad University of Medical Sciences for providing this great opportunity for students like them.

Author statements

Ethical approval

This study did not have any ethical approval as its design (systematic review and meta-analysis) does not require any ethical approval.

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.2021.11.005.

References

1. Leith A, Sharpe JA. Deep-Focus earthquakes and their geological significance. *J Geol* 1936;**44**(8):877–917.
2. Blaikie P, Cannon T, Davis I, Wisner B. *At risk: natural hazards, people's vulnerability and disasters*. Routledge; 2005.
3. Kessler RC, Sonnega A, Bromet E, Hughes M, Nelson CB. Posttraumatic stress disorder in the national comorbidity survey. *Arch Gen Psychiatr* 1995;**52**(12):1048–60.
4. Watson JT, Gayer M, Connolly MA. Epidemics after natural disasters. *Emerg Infect Dis* 2007;**13**(1):1–5.
5. Kouadio IK, Aljunid S, Kamigaki T, Hammad K, Oshitani H. Infectious diseases following natural disasters: prevention and control measures. *Expert Rev Anti-infect Ther* 2012;**10**(1):95–104.
6. Benedict K, Park BJ. Invasive fungal infections after natural disasters. *Emerg Infect Dis* 2014;**20**(3):349–55.
7. Floret N, Viel JF, Mauny F, Hoen B, Piarroux R. Negligible risk for epidemics after geophysical disasters. *Emerg Infect Dis* 2006;**12**(4):543–8.
8. Pourhosseini SS, Ardalan A, Mehrolhassani MH. Key aspects of providing healthcare services in disaster response stage. *Iran J Public Health* 2015;**44**(1):111–8.
9. Jafari N, Shahsanai A, Memarzadeh M, Loghmani A. Prevention of communicable diseases after disaster: a review. *J Res Med Sci : the official journal of Isfahan University of Medical Sciences* 2011;**16**(7):956–62.
10. Ukai T. The great hanshin-awaji earthquake and the problems with emergency medical care. *Ren Fail* 1997;**19**(5):633–45.
11. Tierney KJ. Emergency medical preparedness and response in disasters: the need for interorganizational coordination. *Publ Adm Rev* 1985;**45**:77–84.
12. Peterson J, Welch V, Losos M, Tugwell P. *The Newcastle-Ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses*. Ottawa: Ottawa Hospital Research Institute; 2011. p. 1–12.
13. Akhavan Rezayat A, Dadgar Moghadam M, Ghasemi Nour M, Shirazinia M, Ghodsi H, Rouhbakhsh Zahmatkesh MR, et al. *Association between smoking and non-alcoholic fatty liver disease: a systematic review and meta-analysis*, vol. 6. SAGE open medicine; 2018. 2050312117745223.
14. Aoyagi T, Yamada M, Kunishima H, Tokuda K, Yano H, Ishibashi N, et al. Characteristics of infectious diseases in hospitalized patients during the early phase after the 2011 great East Japan earthquake: pneumonia as a significant reason for hospital care. *Chest* 2013 Feb 1;**143**(2):349–56.
15. Daito H, Suzuki M, Shiihara J, Kilgore PE, Ohtomo H, Morimoto K, et al. Impact of the Tohoku earthquake and tsunami on pneumonia hospitalisations and mortality among adults in northern Miyagi, Japan: a multicentre observational study. *Thorax* 2013;**68**(6):544–50.
16. Fakoorziba M, Baseri A, Eghbal F, Rezaee S, Azizi K, Moemenbellah-Fard M. Post-earthquake outbreak of cutaneous leishmaniasis in a rural region of southern Iran. *Ann Trop Med Parasitol* 2011;**105**(3):217–24.
17. Vasquez D, Palacio A, Nuñez J, Briones W, Beier JC, Pareja DC, et al. Impact of the 2016 Ecuador earthquake on Zika virus cases. *Am J Publ Health* 2017;**107**(7):1137–42.
18. Zhang S, Lu Z, Liu H, Xiao X, Zhao Z, Bao G, et al. Incidence of Japanese encephalitis, visceral leishmaniasis and malaria before and after the Wenchuan earthquake, in China. *Acta Trop* 2013;**128**(1):85–9.
19. 2010 international conference on power system technology. In: Araneda JC, Rudnick H, Mocarquer S, Miquel P, editors. *Lessons from the 2010 Chilean earthquake and its impact on electricity supply*. IEEE; 2010.
20. Dai J, Zhang D, Lu H. Preventive strategies for communicable diseases after earthquakes. *J Sun Yat-sen Univ (Soc Sci Ed)* 2008;**29**(5):506–11.
21. Dube A, Moffatt M, Davison C, Bartels S. Health outcomes for children in Haiti since the 2010 earthquake: a systematic review. *Prehospital Disaster Med* 2018;**33**(1):77–88.
22. Aflatoonian MR, Sharifi I, Poursmaelian S, Hakimi-Parizi M, Ziaali N. The emergence of anthroponotic cutaneous leishmaniasis following the earthquake in southern villages of Bam district, southeastern Iran, 2010. *Journal of Arthropod-Borne Diseases* 2013;**7**(1):8.
23. Sharifi I, Poursmaelian S, Aflatoonian MR, Ardakani RF, Mirzaei M, Fekri AR, et al. Emergence of a new focus of anthroponotic cutaneous leishmaniasis due to Leishmania tropica in rural communities of Bam district after the earthquake, Iran. *Trop Med Int Health* 2011;**16**(4):510–3.
24. Kamigaki T, Seino J, Tohma K, Nukiwa-Soma N, Otani K, Oshitani H. Investigation of an Influenza A (H3N2) outbreak in evacuation centres following the Great East Japan earthquake, 2011. *BMC Publ Health* 2014;**14**(1):34.
25. Kawano T, Hasegawa K, Watase H, Morita H, Yamamura O. Infectious disease frequency among evacuees at shelters after the great eastern Japan earthquake and tsunami: a retrospective study. *Disaster Med Public Health Prep* 2014;**8**(1):58–64.
26. Kawano T, Tsugawa Y, Nishiyama K, Morita H, Yamamura O, Hasegawa K. Shelter crowding and increased incidence of acute respiratory infection in evacuees following the Great Eastern Japan Earthquake and tsunami. *Epidemiol Infect* 2016;**144**(4):787–95.
27. Feng J, Xia Z, Zhang L, Cheng S, Wang R. Risk assessment of malaria prevalence in Ludian, Yongshan, and Jinggu counties, Yunnan province, after 2014 earthquake disaster. *Am J Trop Med Hyg* 2016;**94**(3):674–8.
28. Townes D, Existe A, Boncy J, Magloire R, Vely J-F, Amsalu R, et al. Malaria survey in post-earthquake Haiti—2010. *Am J Trop Med Hyg* 2012;**86**(1):29–31.
29. Adams J. *Environmental health in emergencies and disasters: a practical guide*. World health organization; 2002.
30. Perry RW, Lindell MK, Tierney KJ. *Facing the unexpected: disaster preparedness and response in the United States*. Joseph Henry Press; 2001.
31. Aldrich N, Benson WF. Peer reviewed: disaster preparedness and the chronic disease needs of vulnerable older adults. *Prev Chronic Dis* 2008;**5**(1).
32. Rawls CG, Turnquist MA. Pre-positioning of emergency supplies for disaster response. *Transp Res Part B Methodol* 2010;**44**(4):521–34.
33. Salama P, Spiegel P, Talley L, Waldman R. Lessons learned from complex emergencies over past decade. *Lancet* 2004;**364**(9447):1801–13.
34. Nampala P, Makara A. *Conference on agro-biotechnology, biosafety and seed systems in developing countries*. 2011.
35. Alexander D. Disease epidemiology and earthquake disaster: The example of Southern Italy after the 23 November 1980 earthquake. *Soc Sci Med* 1982 Jan 1;**16**(22):1959–69.
36. Furusawa T, Furusawa H, Eddie R, Tuni M, Pitakaka F, Aswani S. Communicable and non-communicable diseases in the Solomon Islands villages during recovery from a massive earthquake in April 2007. *NZ Med J* 2011 Apr 29;**124**(1333):17–28.
37. Greco D, Faustini A, Forastiere F, Galanti MR, Magliola ME, Moro ML, Piergentili P, Rosmini F, Stazi MA, Luzi S, Fantozzi L. Epidemiological surveillance of diseases following the earthquake of 23rd November 1980 in Southern Italy. *Disasters* 1981 Dec 1;**5**(4):398–406.
38. Jonaidi Jafari N, Radfar MH, Ghofrani H. Incidence of infectious diseases one month after the bam eartquake (2004). *J Med Sci* 2007;**7**(4):597–602.
39. Karmakar S, Rathore AS, Kadri SM, Dutt S, Khare S, Lal S. Post-earthquake outbreak of rotavirus gastroenteritis in Kashmir (India): An epidemiological analysis. *Public Health* 2008 Oct 1;**122**(10):981–9.
40. Khan S, Rai MA, Khan A, Farooqui A, Kazmi SU, Ali SH. Prevalence of HCV and HIV infections in 2005-Earthquake-affected areas of Pakistan. *BMC Infect Dis* 2008 Dec;**8**(1):1–7.
41. Matsuoka T, Yoshioka T, Oda J, Tanaka H, Kuwagata Y, Sugimoto H, Sugimoto T. The impact of a catastrophic earthquake on morbidity rates for various illnesses. *Public Health* 2000 Jul 1;**114**(4):249–53.
42. Ortiz MR, Le NK, Sharma V, Hoare I, Quizhpe E, Teran E, Naik E, Salihu HM, Izurieta R. Post-earthquake Zika virus surge: Disaster and public health threat amid climatic conduciveness. *Sci Rep* 2017 Nov 13;**7**(1):1–0.
43. Salazar MA, Pesigan A, Law R, Winkler V. Post-disaster health impact of natural hazards in the Philippines in 2013. *Global Health Act* 2016 Dec 1;**9**(1):31320.
44. Schneider E, Hajjeh RA, Spiegel RA, Jibson RW, Harp EL, Marshall GA, Gunn RA, McNeil MM, Pinner RW, Baron RC, Burger RC. A coccidioidomycosis outbreak following the Northridge, Calif, earthquake. *JAMA* 1997 Mar 19;**277**(11):904–8.
45. Tohma K, Suzuki A, Otani K, Okamoto M, Nukiwa N, Kamigaki T, Kawamura K, Nakagawa H, Oshitani H. Monitoring of influenza viruses in the aftermath of the Great East Japan earthquake. *Japanese J Infect Dis* 2012;**65**(6):542–4.

Public Health 202 (2022) 32–34

Contents lists available at ScienceDirect

Public Health

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

Short Communication

Is the pandemic leading to a crisis of trust? Insights from an Italian nationwide study

M.R. Gualano, G. Lo Moro*, G. Voglino, F. Bert, R. Siliquini

Department of Public Health Sciences, University of Turin, Turin, Italy

ARTICLE INFO

Article history:
Received 30 July 2021
Received in revised form
15 October 2021
Accepted 29 October 2021
Available online 5 December 2021

Keywords:
COVID-19
Trust
Pandemic fatigue

ABSTRACT

Objectives: Along with mistrust toward politics and journalism, the pandemic is amplifying mistrust in healthcare. To explore trust in key professionals among the Italian population, we focused on perceived change in trust during the pandemic.
Study design: Nationwide online cross-sectional survey (called COCOS).
Methods: COCOS was conducted in Italy in two periods: the end of the first lockdown (T1: April–May 2020) and the end of 2020 (T2: November–December 2020). Descriptive analyses and multivariable logistic regressions were performed (sample size = 2673).
Results: Trust in healthcare workers (HCWs) was reduced in 1.5% of participants (T1) and 2.8% (T2). Trust in scientists/researchers was reduced in 5.8% (T1) and 7.6% (T2). Trust in politicians was reduced in 37.6% (T1) and 52.3% (T2). Trust in journalists was reduced in 41.7% (T1) and 48.3% (T2). Considering multivariable models, participants of the second period, participants who were HCWs, participants with anxiety symptoms, and those experiencing economic struggle due to the pandemic had a higher likelihood of having a reduced trust. The period had the strongest association with reduced trust.
Conclusions: We argue that a central role might be played by the pandemic fatigue. We suggest leading figures should be more aware of the relationship between communication and trust. The pandemic is a real-world experiment in reshaping mediated communication and, although social media play an important role, other approaches might be successful. As a notable part of the population is trusting politicians and media less and less, Italian key professionals should implement initiatives to reinvigorate public support.

Introduction

Already before the pandemic, decreasing trends in trust in governments and media had been identified worldwide, thus causing concerns for potential economic and political impacts.[1] Such crisis of trust grew even more in the context of the pandemic, including in countries where the overall attitude of people toward their government's decisions was positive.[1] Along with mistrust toward politics and journalism, the pandemic is also amplifying mistrust in health care,[2] which might lead to avoidance of health care services[3] and potential damages for public health. To explore trust in key figures among the Italian adult population and its evolution across the pandemic, we focused on perceived change in trust as part of the Covid Collateral Impacts (COCOS) project.[4]

Methods

COCOS was a nationwide online cross-sectional survey conducted in two periods: the end of the first lockdown (April–May 2020)[4] and the end of 2020 (November–December 2020), when restrictions were more relaxed. The Ethics Committee of the University of Torino approved the protocol and written informed consent was obtained from all participants. The online questionnaire was distributed through social networks (mainly Facebook and Twitter) from the institutional pages of the School of Public Health (University of Torino). Participants were recruited by convenience, and participation was voluntary and without compensation. Inclusion criteria were having 18 years or more and living in Italy.

* Corresponding author. Department of Public Health Sciences, University of Turin,Via Santena 5 bis, 10126, Turin, Italy. Tel.: +390116705837; fax: +390116795889.
E-mail address: giuseppina.lomoro@unito.it (G. Lo Moro).

https://doi.org/10.1016/j.puhe.2021.10.015
0033-3506/

The questionnaires used in the two periods had identical items, thus allowing to compare the data. However, the questionnaire used in the second period was shorter and included fewer sections with the aim of gaining more participation. The questionnaire used in the first period of observation is fully described in previous papers from COCOS data.[3,4] The questionnaire distributed in the second period included three sections. The first section was about sociodemographic characteristics, such as age, gender, nationality, marital status, education level, economic situation related to the pandemic, job, and health condition. The second part collected information about opinion and behaviors, e.g. trust and time spent on the Internet. Last, mental health and access to health care during the pandemic were investigated. We used the Generalized Anxiety Disorder 2-item (GAD-2) to screen for symptoms of anxiety, and we used the threshold of $\geq$3 as it provides the highest sensitivity/specificity balance for the GAD-2 in recognizing generalized anxiety disorder.[5]

Specifically, the present short communication focuses on the items about trust. In both periods, participants were asked if their level of trust in healthcare workers (HCWs), scientists/researchers, politicians, technical-scientific consultants of the government (TSCs), and journalists was unchanged/increased/reduced due to the pandemic. We think that the unprecedented situation and the uncertainties that accompany the pandemic may contribute to the rise of mistrust, along with conspiracies, misinformation, and ineffective communication by authorities.[1] Thus, our hypothesis was that trust in all key professionals was reduced in both periods, similar to the crisis of trust recorded in other contexts during the pandemic.[1,2]

Statistical analysis

Descriptive analyses of the level of trust in the above-mentioned professionals were performed, stratifying by the period of observation. Then, for each professional category, trust was dichotomized in reduced versus unchanged/increased (outcome: reduced trust). Multivariable logistic regression models were performed primarily to explore if the period of observation had a significant association with the outcome. In addition to the period of observation, the models were adjusted for: age, gender, education level, being an HCW, experiencing economic struggle due to the pandemic, GAD-2 score$\geq$3, increased time spent on the Internet during the pandemic, having a chronic disease, and having a member of the family who is an HCW. Results were expressed as adjusted Odds Ratio (adjOR) and 95% Confidence Interval (CI).

SPSS (v27) was used, and a two-tailed *P*-value < 0.05 was considered significant. Missing values were excluded.

Results

A total of 2747 individuals answered the questionnaire. However, 74 questionnaires (2.7%) were excluded from all analyses since 70 participants lived outside Italy and 4 participants were aged less than 18 years. Thus, the final sample consisted of 2673 participants (first period = 1515; second period = 1158). Females accounted for 66.8% and the mean age was 44 years (SD = 15). Table 1 shows the main findings. The lowest reductions in trust were reported toward HCWs and scientists/researchers, while the greatest reductions were shown toward politicians and journalists. Specifically, trust in HCWs was reduced in 1.5% of the sample in the first period and in 2.8% in the second; trust in scientists/researchers was reduced in 5.8% in the first period and in 7.6% in the second. Trust in politicians was reduced in 37.6% in the first period and in 52.3% in the second; trust in TSCs was reduced in 23.7% in the first period and in 35.5% in the second; trust in journalists was reduced in 41.7% in the first period and 48.3% in the second. An increased trust was reported by over 27% regarding HCWs and scientists/researchers in both the periods (with a maximum of 43% in the first period for HCWs), while trust in the other categories was increased in less than 10% (with a maximum of 9.3% in the first period for TSCs). On average, across all the categories of professionals, 22.0% (first period) and 29.6% (second period) of people selected 'reduced trust', 59.9% (first period) and 53.3% (second period) selected 'unchanged trust', and 18.1% (first period) and 17.0% (second period) selected 'increased trust'.

In considering multivariable models, participants of the second period were significantly more likely to report a reduced trust in all the professionals, except for scientists/researchers ($P = 0.112$). It is worth noting that participants who were HCWs had a higher likelihood of having a reduced trust in politicians and journalists. Similarly, a reduced trust was more likely among participants with anxiety symptoms (for politicians, journalists, TSCs) and those experiencing economic struggle due to the pandemic (for HCWs, politicians, and TSCs). Most often, the period of observation had the strongest association with reduced trust compared with the other variables.

Discussion

As expected,[1] COCOS highlighted a substantially decreased trust in politics and journalism professionals, while the reduction of trust in purely scientific figures was remarkably lower. Particularly, the decline in trust was significantly higher at the end of 2020. Interestingly, another Italian study showed a steady level of trust toward authorities and considering the beginning of the lockdown and the period right after the lockdown.[6] This could suggest that trust might be not dependent on restrictive measures, but the duration of the pandemic and its long-term correlates may largely contribute to trust fluctuations and the public opinion may change promptly. Accordingly, a scoping review concluded that trust is not stable over time; however, no longitudinal studies explored such variability and the determinants should be further investigated.[7] We argue that a central role might be played by the pandemic fatigue i.e. 'demotivation to follow recommended protective behaviors, emerging gradually over time', which is likely to increase as people go through the personal, social and economic repercussions of restrictions.[8] In light of this, our results about the lower trust perceived by HCWs, people with anxiety and with economic struggle appear clear: citizens who are bearing a high burden due to the pandemic per se and its restrictions might be no longer able to balance and understand the risk of the disease and the measures that must be implemented to tackle COVID-19. Our findings are in line with the conclusions of Barrafrem and colleagues, which showed that the trust in governmental institutions during the pandemic had an important impact on the general well-being and mental health of the population, especially through the mediation of financial well-being.[9]

The present study had some limitations, such as convenience sampling and the cross-sectional design. In addition, considering the online distribution, no data about individuals who refused to participate were registered, and no refusal rate was available. Last, some of the significant CIs of our multivariable models were very close to 1.00 (especially for the relationships with age), thus making it difficult to infer the statistical significance of certain results, which can be interpreted only in a cautious and explorative way. However, to our knowledge, it was one of the first studies exploring trust among the general Italian population in two different pandemic periods.

We suggest that leading figures should be more aware of the relationship between communication and trust. The pandemic has

Table 1
Perceived change in trust in key figures during the pandemic: descriptive analysis and multivariable logistic regression models from the COCOS project (Italy, 2020).

	Trust in healthcare workers		Trust in scientists and researchers		Trust in politicians		Trust in technical-scientific consultants of the government		Trust in journalists	
	1st period n = 1515 %	**2nd period** n = 1158 %	**1st period** n = 1515 %	**2nd period** n = 1158 %	**1st period** n = 1515 %	**2nd period** n = 1158 %	**1st period** n = 1515 %	**2nd period** n = 1158 %	**1st period** n = 1515 %	**2nd period** n = 1158 %
Descriptive analysis										
Unchanged	55.1	54.3	67.2	61.4	54.0	43.8	67.0	57.1	56.3	50.0
Increased	43.4	41.2	27.0	31.0	8.4	3.9	9.3	7.5	2.0	1.7
Reduced	1.5	2.8	5.8	7.6	37.6	52.3	23.7	35.5	41.7	48.3
Multivariable model (outcome: reduced trust)[a]	**adjOR**	**95% CI**	**adjOR**	**95% CI**	**adjOR**	**95% CI**	**adjOR**	**95% CI**	**adjOR**	**95% CI**
Female	1.23	0.69–2.19	**0.61**	**0.44–0.86**[b]	1.00	0.83–1.19	0.91	0.75–1.10	**0.79**	**0.67–0.95**[c]
Age	**1.02**	**1.0003–1.04**[c]	**1.03**	**1.02–1.04**[a]	**1.01**	**1.005–1.02**[b]	**1.02**	**1.01–1.03**[a]	**0.99**	**0.98–0.996**[b]
Second Period	**4.66**	**2.61–8.32**[a]	1.33	0.94–1.88	**1.74**	**1.46–2.08**[a]	**1.68**	**1.39–2.03**[a]	**1.25**	**1.05–1.49**[b]
Having at least a bachelor's degree	**2.13**	**1.21–3.76**[b]	1.00	0.7–1.43	**0.68**	**0.57–0.81**[a]	**0.70**	**0.58–0.84**[a]	1.00	0.84–1.19
Healthcare worker	1.13	0.64–2.02	1.23	0.82–1.85	**1.36**	**1.11–1.67**[b]	1.24	0.99–1.54	**1.33**	**1.09–1.62**[b]
Economic struggle due to the pandemic	**2.53**	**1.51–4.24**[a]	1.39	0.97–1.98	**1.39**	**1.15–1.67**[b]	**1.36**	**1.11–1.66**[b]	1.07	0.89–1.29
GAD-2 score ≥3 (cut-off for the screening of GAD)	1.38	0.78–2.43	1.28	0.87–1.89	**1.48**	**1.21–1.79**[a]	**1.52**	**1.23–1.87**[a]	**1.35**	**1.11–1.63**[b]

n = sample size.
1st period: from April 19th to May 3rd 2020.
2nd period: from November 29th to December 27th 2020.
***P*-value:** $P < 0.001$[a], $P < 0.010$[b], $P < 0.050$[c].
Abbreviations: adjOR adjusted Odds Ratio; CI Confidence Interval; GAD Generalized Anxiety Disorder; GAD-2 Generalized Anxiety Disorder 2-item.
In bold: *P* < 0.050.
[a] The models were also adjusted for increased time spent on the Internet during the pandemic; having a chronic disease; having a member of the family who is a healthcare worker. No significant associations with these variables were revealed.

been a real-world experiment in reshaping mediated communication[1] and, although social media play an important role in political influence, other approaches might be more successful. Teufel and colleagues have interestingly explored the example of Angela Merkel: concurrent with the German Chancellor's speech, they reported a higher level of trust in governmental policies and a reduction of anxiety and depression among the German population.[10] Therefore, as a notable part of the population is trusting politicians and media less and less, Italian key professionals should immediately implement initiatives and policies aiming to reinvigorate public support.

Author statements

Ethical approval

All procedures performed were in accordance with the 1964 Helsinki declaration and its later amendments. The Ethics Committee of the University of Torino approved the protocol (Prot. n. 336581). Written informed consent was obtained from all participants.

Funding

None declared.

Competing interests

None declared.

References

1. Flew T. The global trust deficit disorder: a communications perspective on trust in the time of global pandemics. *J Commun* 2021;**71**(2):163–86. https://doi.org/10.1093/joc/jqab006.
2. Jain SH, Lucey C, Crosson FJ. The enduring importance of trust in the leadership of health care organizations. *JAMA* 2020;**324**(23):2363–4. https://doi.org/10.1001/jama.2020.18555.
3. Gualano MR, Corradi A, Voglino G, Bert F, Siliquini R. Beyond COVID-19: a cross-sectional study in Italy exploring the covid collateral impacts on healthcare services. *Health Pol* 2021;**125**(7):869–76. https://doi.org/10.1016/j.healthpol.2021.03.005.
4. Gualano MR, Lo Moro G, Voglino G, Bert F, Siliquini R. Effects of covid-19 lockdown on mental health and sleep disturbances in Italy. *Int J Environ Res Publ Health* 2020;**17**(13):4779. https://doi.org/10.3390/ijerph17134779.
5. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic meta-analysis. *Gen Hosp Psychiatry* 2016;**39**:24–31. https://doi.org/10.1016/j.genhosppsych.2015.11.005.
6. Graffigna G, Palamenghi L, Savarese M, Castellini G, Barello S. Effects of the COVID-19 emergency and national lockdown on Italian citizens' economic concerns, government trust, and health engagement: evidence from a two-wave panel study. *Milbank Q* 2021;**99**(2):369–92. https://doi.org/10.1111/1468-0009.12506.
7. Berg SH, O'Hara JK, Shortt MT, et al. Health authorities' health risk communication with the public during pandemics: a rapid scoping review. *BMC Publ Health* 2021;**21**(1):1401. https://doi.org/10.1186/s12889-021-11468-3. Published 2021 Jul 15.
8. World Health Organization (WHO). *Pandemic fatigue: reinvigorating the public to prevent covid-19*. 2020. https://apps.who.int/iris/bitstream/handle/10665/335820/WHO-EURO-2020-1160-40906-55390-eng.pdf. [Accessed 8 July 2021].
9. Barrafrem K, Tinghög G, Västfjäll D. Trust in the government increases financial well-being and general well-being during COVID-19. *J Behav Exp Finance* 2021;**31**:100514. https://doi.org/10.1016/j.jbef.2021.100514.
10. Teufel M, Schweda A, Dörrie N, et al. Not all world leaders use Twitter in response to the COVID-19 pandemic: impact of the way of Angela Merkel on psychological distress, behaviour and risk perception. *J Public Health (Oxf)*. 2020;**42**(3):644–6. https://doi.org/10.1093/pubmed/fdaa060.

Public Health 202 (2022) 52–57

Contents lists available at ScienceDirect

Public Health

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

Original Research

Mortality comparisons of COVID-19 with all-cause and non-communicable diseases in Cyprus, Iceland and Malta: lessons learned and forward planning

S. Cuschieri [a, *], E. Pallari [b], A. Hatziyianni [c], R. Sigurvinsdottir [d], I.D. Sigfusdottir [d, e], Á.K. Sigurðardóttir [f, g]

[a] Department of Anatomy, Faculty of Medicine and Surgery, University of Malta, Msida, Malta
[b] University College London, MRC Clinical Trials and Methodology Unit, London, England
[c] Ammochostos General Hospital, 25 Christou Kkeli, Paralimni, 5310, Cyprus
[d] Department of Psychology, Reykjavik University, Reykjavik, Iceland
[e] Teacher's College, Columbia University, New York, NY, USA
[f] School of Health Science, University of Akureyri, Sólborg, Iceland
[g] Akureyri Hospital, Akureyri, Iceland

ARTICLE INFO

Article history:
Received 18 January 2021
Accepted 22 March 2021
Available online 21 April 2021

Keywords:
Coronavirus
Mortality
Case fatality rate
Non-communicable disease
Islands
Syndemic
Cyprus
Iceland
Malta

ABSTRACT

Objectives: The COVID-19 pandemic has spread throughout the world, including Cyprus, Iceland and Malta. Considering the small population sizes of these three island countries, it was anticipated that COVID-19 would be adequately contained and mortality would be low. This study aims to compare and contrast COVID-19 mortality with mortality from all causes and common non-communicable diseases (NCDs) over 8 months between these three islands.

Methods: Data were obtained from the Ministry of Health websites and COVID dashboards from Cyprus, Iceland and Malta. The case-to-fatality ratio (CFR) and years of life lost (YLLs) were calculated. Comparisons were made between the reported cases, deaths, CFR, YLLs, swabbing rates, restrictions and mitigation measures.

Results: Low COVID-19 case numbers and mortality rates were observed during the first wave and transition period in Cyprus, Iceland and Malta. The second wave saw a drastic increase in the number of confirmed cases and mortality rates, especially for Malta, with high CFR and YLLs. Similar restrictions and measures were evident across the three island countries. Results show that COVID-19 mortality was generally lower than mortality from NCDs.

Conclusions: The study highlights that small geographical and population size, along with similar restrictive measures, did not appear to have an advantage against the spread and mortality rate of COVID-19, especially during the second wave. Population density, an ageing population and social behaviours may play a role in the burden of COVID-19. It is recommended that a country-specific syndemic approach is used to deal with the local COVID-19 spread based on the population's characteristics, behaviours and the presence of other pre-existing epidemics.

Introduction

The COVID-19 pandemic spread across the European continent in early 2020, affecting every country in Europe, including the islands of Cyprus, Iceland and Malta. The Republic of Cyprus and Malta are situated within the Mediterranean Sea, while Iceland is situated between the North Atlantic and Arctic oceans. The three islands share similar population characteristics, including a total population of <1 million (Cyprus 875,900; Malta 514,564 and

* Corresponding author.
E-mail addresses: sarah.cuschieri@um.edu.mt (S. Cuschieri), e.pallari@ucl.ac.uk (E. Pallari), hatziyianniamalia@gmail.com (A. Hatziyianni), rannveigs@ru.is (R. Sigurvinsdottir), ingadora@ru.is (I.D. Sigfusdottir), arun@unak.is (Á.K. Sigurðardóttir).

https://doi.org/10.1016/j.puhe.2021.03.025
0033-3506/

Iceland 364,134) and similar life expectancy among males (Cyprus 78.5 years; Malta 78.9 years and Iceland 79.8 years).

Of these three countries, Iceland reported the first COVID-19 case on the 28th February 2020, followed by Malta (7th March 2020) and Cyprus (9th March 2020).[1] Similar to other countries, their governments implemented a number of restrictions to curb the viral spread during the first wave, and these were subsequently reintroduced in late summer when the second wave prevailed.[1] It was hypothesised that the small geographical and population sizes of these three islands would be an advantage, in addition to the implementation of COVID-19 measures, to curb the viral spread and keep the mortality rate low. This study aims to compare and contrast COVID-19 mortality with mortality from all causes and common non-communicable diseases (NCDs) over 8 months (March to November 2020) between the three small islands of Cyprus, Iceland and Malta.

Methods

Data were obtained from the Ministry of Health websites and COVID-19 dashboards from Cyprus, Iceland and Malta, in addition to local published studies. Comparisons were made between the reported cases, deaths, swabbing rates, restrictions and mitigation strategies. No distinction was made between individuals dying 'with' COVID-19 and individuals dying 'due to' COVID-19 as a result of lack of such data from Iceland and Malta.

The reported COVID-19–positive cases and deaths were subdivided into three phases: (1) the first wave (6th March to 7th May); (2) the transitional period (8th May to 13th August) and (3) the second wave (14th August to 30th November). The case-to-fatality ratio (CFR) for each COVID-19 phase was calculated by subdividing the confirmed number of positive cases by the confirmed number of deaths and then multiplying by 100.[2]

Years of life lost (YLLs) is a metric used in population health to measure the number of years lost due to premature death from a particular cause. This metric is calculated by identifying the number of deaths in an age group and multiplying it by a standard life expectancy for that particular age group. The World Health Organization life expectancy age group tables for each country were used for this analysis.[3] The number of deaths by age groups and gender was obtained from the COVID-19 dashboards of each island. The COVID-19 YLLs for each island (Cyprus, Iceland and Malta) were compared with the YLLs of common NCDs (cardiovascular disease, stroke, chronic respiratory disease, diabetes mellitus) and road traffic injuries, as reported by the Global Burden of Disease for the year 2019.[4]

Results

Covid-19 situation in Cyprus, Iceland and Malta

From the onset of the pandemic until the end of November 2020, Cyprus reported 10,583 COVID-19–positive cases (1206 per 100,000), Iceland reported 5027 positive cases (1381 per 100,000) and Malta reported 9877 positive cases (1919 per 100,000).[5–7] Over the study period of 8 months, the daily average number of new COVID-19 cases was 223 for Cyprus, 14 for Iceland and 117 for Malta. During the first wave (March–May 2020), Iceland reported the highest number of positive cases, as shown in Fig. 1. However, of the three islands, Malta entered the second wave earlier and reported the highest number of positive cases.[1,8] In mid-October, Cyprus reported a spike in cases, exceeding the number of positive cases reported by Iceland and Malta (Fig. 1). Concurrently, the polymerase chain reaction (PCR) COVID-19 swabbing test rate increased over time across the three islands (Fig. 2), with Cyprus reaching 70,714.35 per 100,000 PCR swabs by the end of November, while Iceland recorded 107,478.84 per 100,000 and Malta 83,542.54 per 100,000.[5–7]

COVID-19 mortality in Cyprus, Iceland and Malta

A synergistic relationship between the number of COVID-19–positive cases and the mortality rate was expected as a result of the high infectivity rate and morbidity related to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Of these three islands, Malta reported the highest rate of COVID-19 mortality since the onset of the pandemic with 27.31 per 100,000 population, followed by Iceland with 7.14 per 100,000 population and Cyprus with 6.39 per 100,000 population. On subdividing the 8-month study period of the pandemic into three phases, as shown in Fig. 3, it can be seen that the islands experienced a higher mortality rate during the second wave, which coincides with higher community spread and identified cases. All three islands reported the spread of COVID-19 within nursing homes, especially during the second wave.[6,8,9] Indeed, Iceland reported extensive COVID-19 spread among the geriatric department of Landakot hospital, as well as within nursing homes and a rehabilitation centre.[10]

CFR was measured to assess the impact of the pandemic on the mortality of the population. Interestingly, even though the second wave showed a higher mortality rate for all three islands, the CFR for Malta exceeded those of the other islands. Indeed, on assessing mortality from COVID-19 as a fraction of the reported all-cause mortality for 2019 for each country, mortality from COVID-19 represented 3.73% of all-cause mortality in Malta, 1.23% in Iceland and 0.64% in Cyprus.[4]

As shown in Fig. 4, when stratification was performed by age and gender, the highest morality occurrences were seen in men and the elderly (>85 years) across the three islands.

COVID-19 mortality vs mortality from common NCDs and injuries

NCDs have been reported to be responsible for a substantial burden of disease, including increases in years-lived with disability (YLDs) and YLLs.[11] Comparisons were made between COVID-19 and common NCDs mortality (per 100,000 population) and YLLs among the populations of Cyprus, Iceland and Malta, as shown in Table 1. COVID-19 mortality and YLLs were calculated for the study period of 8 months, whereas NCD and injury mortality and YLLs were calculated for the year 2019. However, it is still evident that COVID-19 has resulted in a substantial burden for each of the three island countries, especially Malta. In fact, in Malta, COVID-19 had a higher mortality rate (per 100,000) than road traffic accidents and diabetes mellitus. A similar picture was observed in Iceland; however, in Cyprus, COVID-19 appeared to have a lower disease burden than common NCDs and injuries (until the end of the study period [November 2020]).

COVID-19 mortality and restrictive measures

The first wave of COVID-19 resulted in the implementation of a number of extreme restrictive measures, including closure of airports and ports (Cyprus and Malta); closure of non-essential retail, bars, restaurants, schools and restricting the number of people in one gathering (Cyprus, Iceland and Malta). Cyprus also restricted free movement of the population to only a daily trip.[1] These measures appear to have been effective as the number of COVID-19 cases and mortality rate remained low (Figs. 1 and 3). Indeed, this led to the slow easing of restrictive measures, while keeping the case numbers and mortality low.[1] However, by mid-August, COVID-19 case numbers began to increase again all over Europe, including

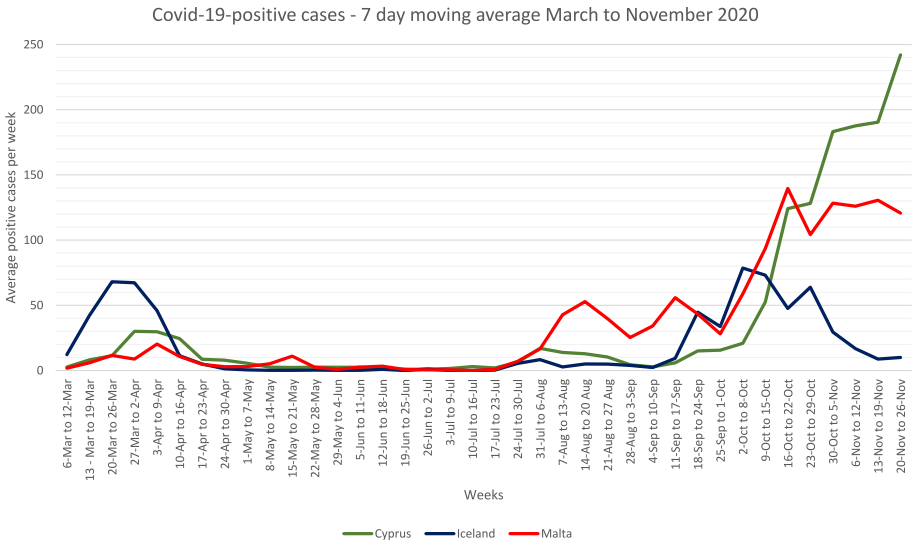


Fig. 1. COVID-19–positive cases: 7-day moving average for Cyprus, Iceland and Malta between March and November 2020.[5–7]

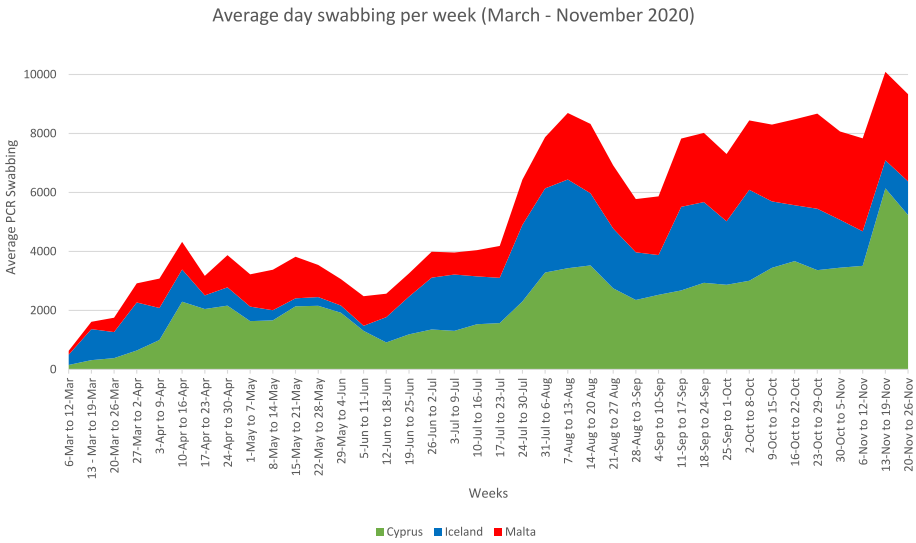


Fig. 2. COVID-19 average PCR swabbing rate for Cyprus, Iceland and Malta between March and November 2020.[5–7] PCR, polymerase chain reaction.

across the three islands. A distinct spike in COVID-19–positive cases was observed in Malta compared with Cyprus and Iceland, especially after the reopening of schools towards the end of September and beginning of October, with a consequential high mortality rate (Fig. 3). A stepwise control restriction was reintroduced, including a reduction in the number of people attending gatherings, compulsory mask wearing indoors and outdoors and restrictions in non-essential services. Supplement Table S1 provides a list of the different restrictions and measures implemented at the onset of the second wave by Cyprus, Iceland and Malta. Although restrictions were put in place, a substantial mortality rate was still evident, especially for Malta and later on for Cyprus.

Discussion

Small islands are considered to be at an advantage when controlling infectious diseases owing to their small population and geographical sizes. In small islands, containment measures at a population level are anticipated to be easier to implement, thus more efficiently limiting the viral spread than larger countries. The absence of land borders further enabled the successful implementation of containment measures. These advantages played a role during the first wave of COVID-19 and the transition period, where Cyprus, Iceland and Malta had relatively low numbers of COVID-19–positive cases and subsequently low mortality rates when compared with larger neighbouring countries.[1,12–14] Of these three small islands, Malta was predominantly in a better containment scenario (not mortality), albeit with their similar restrictions and measures.[1,15]

However, hasty relaxation of restrictions and opening of the airport and ports, in addition to the organisation of large mass events led to the downfall of the stable COVID-19 situation in Malta, giving rise to the second wave.[8] Cyprus and Iceland also experienced an increase in COVID-19–positive cases around late summer time. However, a steeper incline in cases along with spikes in mortality occurred in Malta. This study showed that, similar to other countries, these small islands showed a higher mortality among the elderly population, particularly among men.[16] Despite

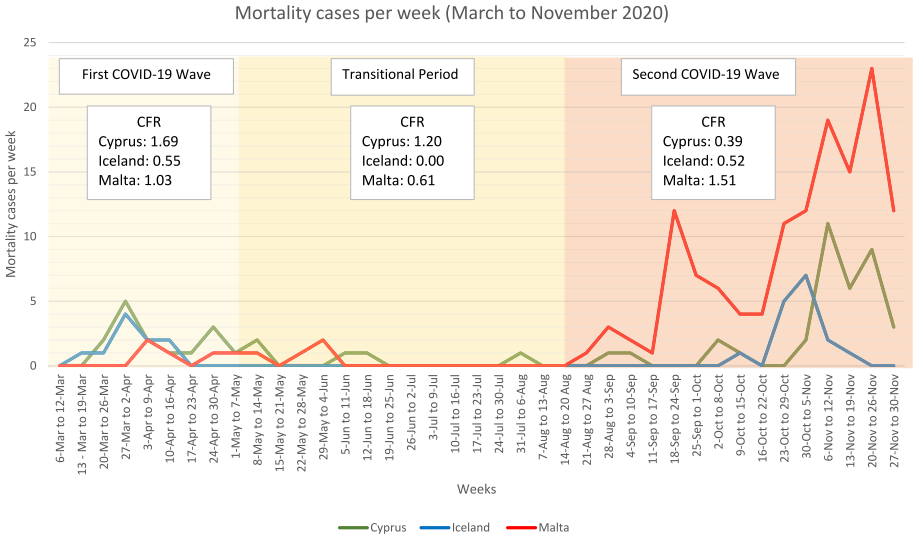


Fig. 3. COVID-19 mortality cases and CFR in Cyprus, Iceland and Malta between March and November 2020. CFR, case-to-fatality ratio.

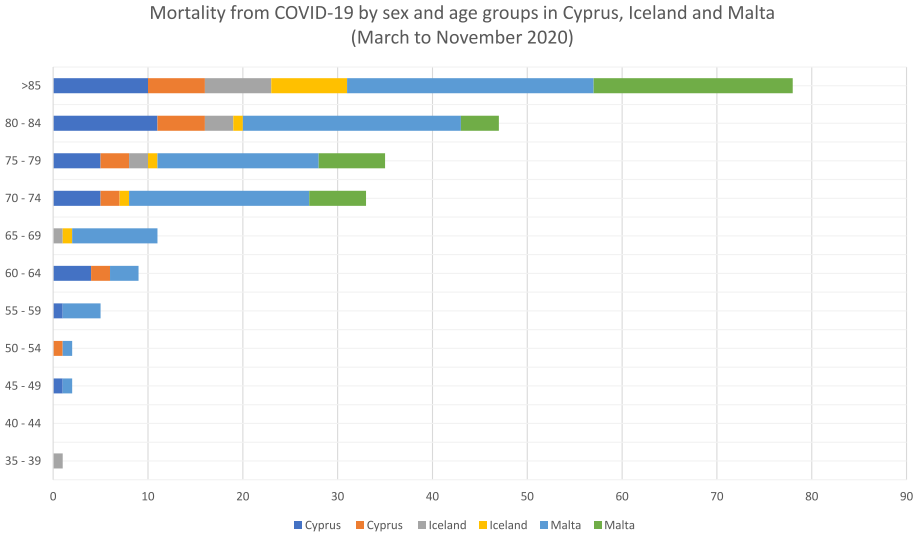


Fig. 4. COVID-19 mortality by age groups and gender in Cyprus, Iceland and Malta between March and November 2020.

Table 1
Comparison between COVID-19 and common non-communicable disease mortality (per 100,000) and YLLs in Cyprus, Iceland and Malta.[a]

Mortality causes	Cyprus		Iceland		Malta	
	Mortality per 100,000	YLLs	Mortality per 100,000	YLLs	Mortality per 100,000	YLLs
COVID-19[b]	6.39	684.33	7.14	273.40	27.40	1558.80
Cardiovascular disease	355.34	48,033.21	198.75	9200.09	289.77	20,557.96
Stroke	85.11	10,176.77	40.76	1774.29	63.96	4281.98
Chronic respiratory disease	61.78	7417.58	30.43	1502.00	30.56	2282.47
Diabetes mellitus	56.00	7086.12	7.13	392.40	23.98	1889.28
Road traffic injuries	15.27	5705.54	3.11	486.51	2.89	649.42

YLLs, years of life lost.
[a] Mortality per 100,000 and YLLs based on the Global Burden of Disease Study 2019 (excl. COVID-19).
[b] Mortality over 8 months (March to November 2020).

similar reintroduction of restrictions at the onset of the second wave in the three islands, the COVID-19 burden in Malta predominated with a higher CFR than that in Cyprus and Iceland. Such CFR differences have been attributed to differences in patient characteristics, prevalence of diagnostic testing and healthcare system availability.[17] However, Cyprus and Malta have been reported to have similar patient metabolic characteristics and healthcare system preparedness for COVID-19.[1,18–19]

The PCR testing strategy in Cyprus, Iceland and Malta has been very similar.[1] However, Iceland reported the highest prevalence of

testing rate, followed by Malta and Cyprus. Furthermore, a spread within nursing homes was reported in all islands,[6,8,9] which has been anticipated to have detrimental impacts on these elderly populations with subsequential high mortality.[20] Indeed, the majority of COVID-19 mortality was among the elderly population in Cyprus, Iceland and Malta.

Considering the similar COVID-19 public health approaches, small populations and geographical aspects shared by Cyprus, Iceland and Malta, the results of this study indicate that other predisposing factors might be contributing to the discrepancy in mortality rates in Malta. Potential factors are a higher ageing population in Malta than Cyprus and Iceland,[4,21] thus increasing susceptibility to COVID-19 morbidity and mortality.[22] Another potential factor could be the social behaviours and attitudes of the populations. The population density of each country may have had an effect on the containment measures and viral spread. Malta's population density is much higher (1380 individuals per km^2) than that of Cyprus (3.31 individuals per km^2) and Iceland (131 individuals per km^2). In fact, countries with low population densities, such as Finland and Norway, have exhibited superior viral containment, which has also been influenced by the timely implementation of measures and compliance and trust within the population.[23]

Notwithstanding the negative COVID-19 pandemic implications on the mortality rate, mental health and well-being, along with the burden on the healthcare systems and economy, the COVID-19 mortality burden (for 8 months) only contributed to a small proportion of the all-cause mortality in Cyprus, Iceland and Malta. Furthermore, the COVID-19 mortality burden was substantially lower than mortality from common NCDs and injuries (with some exceptions). Of note, individuals with NCDs who had COVID-19 were more likely to experience severe infection, leading to higher morbidity and mortality.[24] However, it appeared that the burden of COVID-19 mortality did not exceed the typical annual NCDs burden. This highlights the importance of implementing a syndemic approach, where healthcare systems and policies target both COVID-19 and NCDs simultaneously.[25]

A number of limitations need to be acknowledged for this study. COVID-19 data and analyses were limited to the available data from online sources, such as dashboards, platforms and other ministerial sites. Mortality data identifying individuals dying 'with' COVID-19 as opposed to dying 'due to' COVID-19 were available for Cyprus, but not for Iceland and Malta. Furthermore, it is expected that a proportion of COVID-19 mortality has been unreported, as individuals may die without officially being identified as COVID-19 positive. This may impact the analyses and conclusions of this study. Another limitation is the lack of data on confounding factors that might have had an impact on mortality rates. All-cause mortality, mortality attributed to NCDs and YLLs were reported for the year 2019, while the COVID-19 mortality and YLLs were calculated based on a period of 8 months. These comparisons provide a general overview of the effect of COVID-19 within the study period. However, conclusive comparisons could not be achieved, especially as the pandemic is ongoing and variables, including the rate of spread, restriction measures and epidemiological data, rapidly change.

Conclusion

The spread of the COVID-19 pandemic exhibits no boundaries and has affected all countries across the world, including the small islands of Cyprus, Iceland and Malta. The present study highlights that small geographical and population sizes, along with similar restrictive public health measures, did not have an advantage against the viral spread and mortality rate, especially during the second wave in Malta. Population density, an ageing population and social behaviours may play a role in the burden of COVID-19. However, the COVID-19 burden appeared to have less impact than common NCDs and injuries on the population. It is therefore recommended that a country-specific syndemic approach should be implemented to mitigate the local spread of COVID-19, taking into account population characteristics, behaviours and other pre-existing epidemics.

Author statements

Acknowledgements

The authors would like to thank Ms Margarita Kyriakou, Press Officer of the Ministry of Health, for the provision and validation of the provided data for Cyprus.

Ethical approval

Not required.

Funding

None declared.

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.2021.03.025.

References

1. Cuschieri S, Pallari E, Hatziyiann A, Sigurvinsdottir R, Sigfusdottir ID, Sigurðardóttir ÁK. Dealing with COVID-19 in small European island states: Cyprus, Iceland and Malta. *Early Hum Dev* 2020 Nov 12:105261 [Internet]. [cited 2020 Nov 22]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/33213965.
2. World Health Organization (WHO). Estimating mortality from COVID-19. *Sci Brief* 2020.
3. Global Health Observatory data repository. *Life tables by country* [Internet]. World Health Organization (WHO). World Health Organization; 2018 [cited 2020 Dec 12]. Available from: https://apps.who.int/gho/data/view.main.LT62060?lang=en.
4. Institute for Health Metrics and Evaluation (IHME). *Global health data exchange*. 2019. Seattle.
5. University of Cyprus. *COVID-19 spread in cyprus* [Internet]. University of Cyprus Research and Innovation Center of Excellence; 2020 [cited 2020 Sep 22]. Available from: https://covid19.ucy.ac.cy/.
6. The Directorate of Health. *The department of civil protection and emergency management. COVID-19 in Iceland* [Internet]. 2020 [cited 2020 Sep 23]. Available from: https://www.covid.is/data.
7. COVID-19 Public Health Response Team - Ministry for Health. *COVID-19 data management system*. 2020.
8. Cuschieri S, Balzan M, Gauci C, Aguis S, Grech V. Mass events trigger Malta's second peak after initial successful pandemic suppression. *J Community Health* 2020 Sep 16:1–8 [Internet]. [cited 2020 Sep 17]. Available from: http://link.springer.com/10.1007/s10900-020-00925-6.
9. Financial Mirror. *COVID19: 35 people test positive at Limassol nursing home - Financial Mirror* [Internet]. Financial Mirror; 2020 [cited 2020 Dec 11]. Available from: https://www.financialmirror.com/2020/11/02/covid19-35-people-test-positive-at-limassol-nursing-home/.
10. Hafstao V. *Cluster of COVID-19 infections puts landspítali on high alert - Iceland Monitor* [Internet]. Iceland Monitor; 2020 [cited 2020 Dec 30]. Available from: https://icelandmonitor.mbl.is/news/news/2020/10/26/cluster_of_covid_19_infections_puts_landspitali_on_/.
11. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 2020 Oct 17;**396**(10258):1204–22 [Internet]. [cited 2020 Nov 22].

Available from: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30925-9/fulltext.
12. Remuzzi A, Remuzzi G. COVID-19 and Italy: what next? *Lancet (London, England)* 2020 Apr 11;**395**(10231):1225–8 [Internet]. [cited 2020 Jul 25]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/32178769.
13. Nurchis MC, Pascucci D, Sapienza M, Villani L, D'Ambrosio F, Castrini F, et al. Impact of the burden of COVID-19 in Italy: results of disability-adjusted life years (DALYs) and productivity loss. *Int J Environ Res Public Health* 2020;**17**(12) [Internet]. [cited 2020 Sep 18]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/32545827.
14. Flynn D, Moloney E, Bhattarai N, Scott J, Breckons M, Avery L, et al. COVID-19 pandemic in the United Kingdom. *Health Policy Technol* 2020 Dec 1;**9**(4): 673–91 [Internet]. [cited 2020 Dec 14]. Available from: https://www.sciencedirect.com/science/article/pii/S2211883720300757.
15. Cuschieri S. COVID-19 panic, solidarity and equity—the Malta exemplary experience. *J Public Health (Bangkok)* 2020 May 30:1–6 [Internet]. [cited 2020 Jun 3]. Available from: http://link.springer.com/10.1007/s10389-020-01308-w.
16. Islam N, Khunti K, Dambha-Miller H, Kawachi I, Marmot M. COVID-19 mortality: a complex interplay of sex, gender and ethnicity. *Eur J Public Health* 2020 Oct 1;**30**(5):847–8 [Internet]. [cited 2020 Oct 10]. Available from: https://academic.oup.com/eurpub/article/30/5/847/5879989.
17. Kang S-J, Jung SI. Age-related morbidity and mortality among patients with COVID-19. *Infect Chemother* 2020 Jun 1;**52**(2):154 [Internet]. [cited 2020 Dec 14]. Available from: https://icjournal.org/DOIx.php?id=10.3947/ic.2020.52.2.154.
18. Cuschieri S, Falzon C, Janulova L, Aguis S, Busuttil W, Psaila N, et al. Malta's only acute public hospital service during COVID-19: a diary of events from the first wave to transition phase. *Int J Qual Health Care* 2021;**33**(1). mzaa138, https://doi.org/10.1093/intqhc/mzaa138.
19. Cuschieri S, Pallari E, Terzic N, Alkerwi A, Sigurðardóttir ÁK. Mapping the burden of diabetes in five small countries in Europe and setting the agenda for health policy and strategic action. *Health Res Policy Sys* 2021;**19**:43. https://doi.org/10.1186/s12961-020-00665-y.
20. Signorelli C, Odone A. Age-specific COVID-19 case-fatality rate: no evidence of changes over time. *Int J Public Health*. [cited 2020 Dec 14];65: 1435–6. Available from: https://doi.org/10.1007/s00038-020-01486-0.
21. Institute for Health Metrics and Evaluation (IHME). *Malta profile* [Internet]. Seattle, WA: IHME, University of Washington; 2019 [cited 2020 Feb 15]. Available from: http://www.healthdata.org/malta.
22. De Larochelambert Q, Marc A, Antero J, Le Bourg E, Toussaint J-F. Covid-19 mortality: a matter of vulnerability among nations facing limited margins of adaptation. *Front Public Health* 2020 Nov 19;**8**:782 [Internet]. [cited 2020 Dec 14]. Available from: https://www.frontiersin.org/articles/10.3389/fpubh.2020.604339/full.
23. Organisation for Economic Co-operation and Development (OECD). *Health at a glance: Europe 2020 state of health in the EU cycle*. 2020 [Internet]. [cited 2020 Dec 15]. Available from: https://doi.org/10.1787/82129230-en.
24. Cuschieri S, Grech V. At risk population for COVID-19: Multimorbidity characteristics of European Small Island State. *Public Health* 2021;**192**:33–6 . https://doi.org/10.1016/j.puhe.2020.12.012.
25. Horton R. Offline: COVID-19 is not a pandemic (London, England) [Internet] *Lancet* 2020 Sep 26;**396**(10255):874. Available from: http://www.ncbi.nlm.nih.gov/pubmed/32979964. [Accessed 14 November 2020].

Public Health 202 (2022) 12–17

Contents lists available at ScienceDirect

Public Health

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

Original Research

NFI, a clinical scoring tool for predicting non-alcoholic fatty liver in the Chinese population

M. Zhao [a, b], Y. Hu [c], C. Shi [d], K. Wang [e], J. Li [c], J. Song [c], C. Huo [c], Y. Xi [b], S. Bu [b], Q. Huang [c, *]

[a] Department of Medical Services, The Affiliated Hospital of Medical School of Ningbo University, 315000, Ningbo, China
[b] Diabetes Research Center, School of Medicine, Ningbo University, 315211 Ningbo, China
[c] Department of Endocrinology, Changhai Hospital, Second Military Medical University, 200433 Shanghai, China
[d] Cixi Institute of Biomedical Engineering, Ningbo Institute of Industrial Technology, Chinese Academy of Sciences, 315201 Ningbo, China
[e] College of Science and Technology, Ningbo University, 315211 Ningbo, China

ARTICLE INFO

Article history:
Received 15 May 2021
Received in revised form
23 September 2021
Accepted 26 October 2021
Available online 4 December 2021

Keywords:
Non-alcoholic fatty liver disease
Predictive model
Validation

ABSTRACT

Objectives: Accurate assessment of early non-alcoholic fatty liver disease (NAFLD) is important to reduce the possible complications. The purpose of the present study was to develop a simple algorithm for the screening of NAFLD in the Chinese population based on routine anthropometric data and laboratory tests.
Study design: This is a cross-sectional design.
Methods: The subjects (1145) underwent routine physical examinations. The variables in the NAFLD index (NFI) were obtained by a stepwise multiple logistic regression analysis on 1000 bootstrap samples. The area under the receiver-operating characteristic (AUROC) was used to evaluate the accuracy of the NFI.
Results: Multivariate analysis showed that body mass index, fasting blood glucose, ratio of alanine aminotransferase to aspartate aminotransferase, and triglyceride were included in the final equation. The AUROC of the NFI was 0.919 (95% confidence interval = 0.901–0.937). An NFI of <31.0 excluded the possibility of NAFLD with a sensitivity of 96.9%, and at a value of >36.0, the NFI could detect NAFLD with a specificity of 98.9%.
Conclusions: NFI was a cost-effective NAFLD-screening model, which had a high accuracy for predicting NAFLD at early stages in the Chinese population.

Introduction

Non-alcoholic fatty liver disease (NAFLD) is a chronic condition characterized by hepatic steatosis in the absence of excessive alcohol use or other specific damage factors.[1] It represented a spectrum of chronic liver disease ranging from simple fatty liver (steatosis) to non-alcoholic steatohepatitis (NASH) and even cirrhosis.[2,3] Many studies suggest that NAFLD is a result of improved living standards, increased dietary fat, and unhealthy lifestyle and that the increased prevalence of NAFLD is in parallel with that of excessive body mass index (BMI), visceral obesity, insulin resistance (IR), type 2 diabetes mellitus, and metabolic syndrome (MetS).[4–9] Obesity is a major public health problem around the world.[10,11] Zhang et al.[12] found that the prevalence of being overweight and obese was 15.2% and 11.7%, respectively, in Jiangsu Province, China. Their results were similar to other studies in China,[13,14] and those results suggested that obesity was a social problem that could not be ignored. Additionally, obesity is a risk factor of NAFLD. A recent meta-analysis indicated that the prevalence of NAFLD in children from general population studies and from studies based on child obesity clinics was 7.6% (95% confidence interval [CI] = 5.5–10.3%) and 34.2% (95% CI = 27.8–41.2%), respectively.[15] Therefore, it was necessary to develop a prediction model to diagnose NAFLD at an early stage in young people.

A large amount of evidence has indicated that NAFLD can lead to severe liver damage in a short period of time; non-alcoholic steatohepatitis may develop in 30% of patients with NAFLD, fibrosis in approximately 25%, cirrhosis in 10–20%, and hepatocellular carcinoma in 4%.[16,17] According to statistics, the global prevalence of NAFLD is 25%, with the highest prevalence in the Middle East and South America and the fibrosis progression proportion and the mean annual rate of progression in NASH being 40.76% and 9%. The

* Corresponding author. Tel.: +86 21 31161392; fax: +86 21 65385289.
E-mail addresses: ming8260@126.com (M. Zhao), 547237104@qq.com (Y. Hu), changchengshi@nimte.ac.cn (C. Shi), wangkaijunkk@outlook.com (K. Wang), lijh0409@163.com (J. Li), 1304482005@qq.com (J. Song), 18221827370@163.com (C. Huo), xiyang@nbu.edu.cn (Y. Xi), bushizhong@nbu.edu.cn (S. Bu), qxinyi1220@163.com (Q. Huang).

https://doi.org/10.1016/j.puhe.2021.10.012
0033-3506/

prevalences of NAFLD in the United States and Europe are 30% and 28%, respectively.[18] In China, the prevalence of NAFLD has been estimated to be approximately 20% in the general adult population.[19] In parts of China with more obese people, the prevalence of NAFLD is up to 51%.[20] Plenty of evidence has indicated that NAFLD can aggravate the progression of cardiovascular diseases and chronic kidney disease.[21–23] In addition, the high prevalence of NAFLD has resulted in an increased incidence of NAFLD-related complications, all-cause mortality, and healthcare costs and declining health-related quality of life.[24]

Most individuals with NAFLD do not have specific symptoms, especially at the early stage, which limits prevention and detection of NAFLD.[25] Liver biopsy is regarded as the golden standard for diagnosing NAFLD, but it is an invasive procedure; furthermore, it only samples a small portion of the liver, which may result in sampling error.[26,27] Imaging modalities such as ultrasonography (US), computed tomography (CT), and magnetic resonance imaging (MRI) have been used as alternatives to screen early-stage NAFLD patients. However, these methods are also not cost-effective, and most early-stage NAFLD patients are virtually asymptomatic. Therefore, several non-invasive scoring systems based on simple clinical or laboratory indices have been developed to identify early-stage fibrosis in patients with NAFLD and other liver diseases. The fatty liver index (FLI) was established by Bedogni et al.[28] and is an algorithm based on BMI, waist circumference (WC), triglycerides (TG), and gamma-glutamyl-transferase (GGT) for the prediction of fatty liver in the general population. The FLI has a high accuracy and has been validated in several different populations.[25,29–31] Other models, for instance, hepatic steatosis index (HIS),[32] BARD score,[33] and ZJU index,[34] are also useful for screening NAFLD. However, because most of these predictive models are based on non-Asians, some parameters may not be suitable for the Chinese population.

Therefore, the purpose of the present study was to develop a simple algorithm for the screening of NAFLD based on routine anthropometric data and laboratory tests and to compare the accuracy of this NAFLD-screening model with previously published models.

Methods

Study participants

The participants were recruited from individuals who had a routine physical examination at the Health Care Center of Changhai Hospital, Shanghai, between 1 October 2012 and 30 November 2012. Participants with the following situations were excluded: excessive alcohol use (>140 g/week in males and >70 g/week in females), presence of hepatitis B surface antigen (HBsAg) or anti-hepatitis C virus antibody (anti-HCV), medications known to precipitate fatty liver (such as amiodarone or tamoxifen) during the previous 6 months, and liver diseases caused due to other reasons. A total of 1145 participants aged between 21 and 83 years were included in the final analysis.

This study was approved by the Ethics Committee of Changhai Hospital. All procedures were in accordance with the ethical standards of the Helsinki Declaration and all participants provided written informed consent.

Anthropometric and clinical measurements

Anthropometric measurements included height and weight. The subjects were requested to empty their bladder, stand upright, and look straight ahead in barefoot and minimal clothing. The height and weight were measured to the nearest millimeter and kilogram, respectively. BMI was calculated as the ratio of weight to the square of height (kg/m^2). Each measurement was completed by two trained investigators, with one investigator taking the measurement and the other recording the reading. Blood pressure was measured three times with 5-minute intervals in a sitting position. Systolic blood pressure (SBP) and diastolic blood pressure (DBP) were measured and the mean values from the three measurements were used for further analysis.

All of the participants were requested to fast overnight for at least 10 h, and then venous blood was collected for biochemical tests. Fasting blood glucose (FBG), alanine aminotransferase (ALT), aspartate aminotransferase (AST), GGT, and serum uric acid (UA) were measured by standard bioassay methods using an automatic biochemistry analyzer (Roche Modular E170, Basel, Switzerland). Total cholesterol (TC), triglyceride (TG), high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C) were assayed enzymatically using a Roche MODULAR DPP automatic biochemistry analyzer. The levels of fasting insulin (FINS) and fasting C peptide (FCP) were determined by radioimmunoassay (Jiuding Biological Engineering Co., Tianjin, China). The following equation was used for the homeostasis model assessment of insulin resistance (HOMA-IR): HOMA-IR = [FINS (uUI/ml) × FPG (mmol/L)]/22.5.[35]

Ultrasonography (US) diagnostic criteria for NAFLD

The hepatic US method was used to diagnose NAFLD according to the following criteria: diffuse increased echogenicity of liver tissues in the near field as compared to the kidney and spleen, and gradual attenuation of liver echogenicity in the far-field; intrahepatic duct structure not clear; enlarged liver with blunted edge angle; and the echo of the right hepatic lobe envelope and the diaphragmatic display not clear or incomplete.[36,37]

Statistical analysis

Continuous variables were tested for normality and presented as means ± standard deviations if normally distributed or median (interquartile range) if they were skewed distributions. Categorical variables were presented as proportions. For the univariate analysis, Student's *t*-test or the Mann–Whitney U-test was used to compare the continuous variables, and the Chi-squared test or Fisher's exact test was used to compare the categorical variables.

Variables that were statistically significant with the area under the receiver-operating characteristic (AUROC) > 0.6 in the univariate analysis were added to the multiple logistic regression on 1000 bootstrap samples (the probabilities to enter was 0.05 and the probabilities to remove was 0.10) to identify the predictors of the presence of NAFLD. Before performing the multiple logistic regression, to obtain a linear logit, some candidate variables underwent a natural logarithm ($\log_e$) transformation. Based on the final result of the multiple logistic regression, a predictive model for NAFLD was established. The goodness of fit of the model was evaluated using the Hosmer–Lemeshow statistic and the accuracy was assessed by the AUROC, sensitivity, specificity, positive likelihood ratio (LR_+), negative likelihood ratio (LR_-), positive predictive values (PPV), and negative predictive values (NPV).

All analyses were performed using SPSS version 19.0 (Chicago, IL, USA) and R version 12.0 (College Station, Texas, USA). Two-tailed tests were used and a *P*-value < 0.05 was considered statistically significant.

Results

Characteristics of the study participants

Among the 1145 subjects who were enrolled in this study, 549 (47.95%) were male, and their mean age (standard deviation) was

44.53 (12.69) years. The NAFLD detection rate by ultrasonography examination was 48.03% of all participants. Among them, male subjects had a significantly higher rate of NAFLD detection rate than female subjects (62.66% vs 34.56%, $\chi^2 = 90.369$, $P < 0.001$). The detection rate of NAFLD in both male and female had a tendency to increase with age (Fig. 1). The demographic and laboratory features of all subjects are provided in Table 1. According to the univariate analysis, in comparison with subjects with no NAFLD, patients with NAFLD were older ($P < 0.001$), were more likely to be male ($P < 0.001$), and had higher SBP and DBP, higher levels of FBG, ALT, AST, GGT, UA, TC, TG, LDL, FINS, and FCP, but significantly lower HDL levels.

Derivation of a new index for NAFLD

The univariate analysis revealed that all of the variables were significantly different between the case and control groups. Subsequently, these variables were included in a multiple logistic regression analysis. However, significant interactions and multicollinearities were detected between FBG, FINS, HOMA-IR, and FCP; between ALT, AST, and ALT/AST; and between TC, TG, HDL-C, and LDL-C. To avoid these interactions and multicollinearities, we incorporated the variables with the highest odds ratios (OR) into the further stepwise logistic regression. Finally, the results suggested that BMI, FBG, ALT/AST, and TG were the independent factors of the presence of NAFLD after adjusting the interactions and multicollinearities between variables (Table 2).

In the multiple logistic regression model, we used these four predictors to calculate the probabilities of NAFLD. It was calculated as $(e^{0.731*BMI+0.599*FBG+2.393*TG+1.052*ALT/AST - 24.218})/(1+e^{0.731*BMI+0.599*FBG+2.393*TG+1.052*ALT/AST - 24.218})$. To simplify this equation, we utilized the exponents and changed the multiplicative factors into approximate integers. In addition, to adjust the difference of BMI between male and female subjects, we added two points to female. As a result, we derived a simple equation, which we named the NFI:

$$\text{NFI} = \text{BMI (kg/m}^2) + \text{FBG (mmol/L)} + 3 \times \text{TG (mmol/L)} + \text{ALT (IU/L)/AST (IU/L)} \ (+2, \text{ if female})$$

Comparison of diagnostic accuracy

To compare the diagnostic accuracy of the NFI and previous models, ROC curves were generated (Fig. 2). The NFI had the best diagnostic accuracy for NAFLD, with the AUROC at 0.919 (95% CI = 0.901–0.937), followed by the ZJU index (AUROC: 0.908, 95% CI = 0.889–0.926), HSI (AUROC: 0.862, 95% CI = 0.840–0.884), and BARD score (AUROC: 0.722, 95% CI = 0.693–0.751). The cut-off point, sensitivity, specificity, LR_+, and LR_- of each test are shown in Table 3. The BARD score had the highest sensitivity of 86.0% and NFI had the highest specificity of 97.6%.

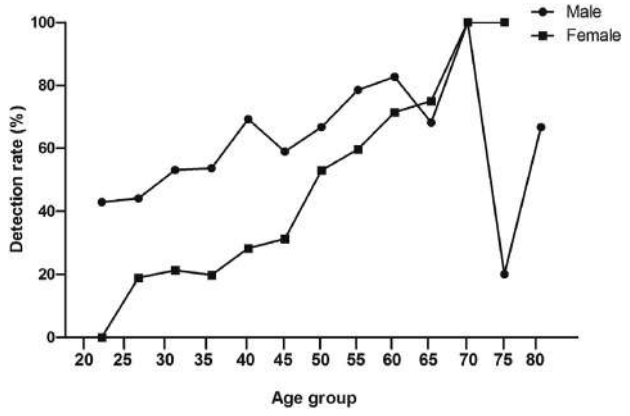


Fig. 1. The NAFLD detection rate of all participants stratified by age and gender.

Table 1
Characteristics of the study participants with and without non-alcoholic fatty liver disease (NAFLD).

Variables	Total	No NAFLD	NAFLD	*P*
Gender (male, %)	549 (47.95)	205 (34.45)	344 (62.55)	<0.001
Age (years)	44.53 ± 12.69	40.68 ± 11.74	48.69 ± 12.37	<0.001
BMI (kg/m^2)	23.96 ± 2.86	22.19 ± 1.29	25.87 ± 2.87	<0.001
SBP (mm Hg)	121.04 ± 17.40	113.85 ± 13.21	128.83 ± 18.03	<0.001
DBP (mm Hg)	73.29 ± 11.47	68.75 ± 8.91	78.19 ± 11.90	<0.001
FBG (mmol/L)	5.34 ± 0.80	5.09 ± 0.40	5.60 ± 1.01	<0.001
ALT (U/L)	19.40 ± 8.77	15.84 ± 7.81	23.24 ± 8.11	<0.001
AST (U/L)	19.52 ± 5.12	18.21 ± 5.08	20.93 ± 4.77	<0.001
ALT/AST ratio	0.98 ± 0.33	0.87 ± 0.31	1.11 ± 0.32	<0.001
GGT (U/L)	23.92 ± 19.11	16.08 ± 9.17	32.41 ± 23.04	<0.001
UA (umol/L)	317.84 ± 84.70	272.64 ± 52.07	366.75 ± 86.05	<0.001
TC (mmol/L)	4.54 ± 0.78	4.23 ± 0.52	4.88 ± 0.87	<0.001
TG (mmol/L)	1.37 ± 0.98	0.90 ± 0.28	1.89 ± 1.18	<0.001
HDL (mmol/L)	1.32 ± 0.30	1.45 ± 0.25	1.18 ± 0.28	<0.001
LDL (mmol/L)	2.82 ± 0.73	2.51 ± 0.48	3.14 ± 0.80	<0.001
FINS (uIU/mL)	18.87 ± 9.34	13.71 ± 6.60	24.45 ± 8.64	<0.001
FCP (ng/ml)	2.84 ± 1.33	2.17 ± 0.92	3.58 ± 1.33	<0.001
HOMA-IR	4.57 ± 2.65	3.10 ± 1.52	6.16 ± 2.69	<0.001

Abbreviations: BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; FBG, fasting blood glucose; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, gamma-glutamyl-transferase; UA, uric acid; TC, total cholesterol; TG, triglyceride; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; FINS, fasting insulin; FCP, fasting C peptide; HOMA-IR, homeostasis model assessment of insulin resistance.

Validation of the NFI

The median value of the NFI in the participants was 33.86 (range 25.42–58.49). For males and females, the AUROCs were 0.960 (95% CI = 0.944–0.976) and 0.889 (95% CI = 0.855–0.923), respectively. An NFI of <31.0 excluded the possibility of NAFLD with a sensitivity of 96.9% (95% CI = 95.0–98.1), and an NPV of 89.9% (84.0–93.8). At a value of >36.0, the NFI could detect NAFLD with a specificity of 98.9% (97.7–99.6) and a positive predictive value of 98.6% (96.7–99.4) (Table 4). Using these cut-off values, 168 subjects (14.7%) had no NAFLD (NFI <31.0) and 416 subjects (36.3%) had NAFLD (NFI >36.0), while 561 subjects (49%) could not be classified ($31.0 \leq \text{NFI} \leq 36.0$).

Discussion

The present study constructed a predictive model of NAFLD by using BMI and the results of routinely performed laboratory tests. In addition, we also validated the performance via comparing it with some previous models.

Large numbers of patients with NAFLD were asymptomatic or had only upper abdominal pain or fatigue before cirrhosis. Only

Table 2
Variables included in the non-alcoholic fatter liver index.

Variables	Odds ratio (OR)	95% CI for OR	*P*
BMI	2.160	1.906–2.447	<0.001
FBG	1.938	1.329–2.825	<0.001
ALT/AST	2.810	1.570–5.030	<0.001
TG	12.024	7.257–19.924	<0.001

Abbreviations: BMI, body mass index; FBG, fasting blood glucose; ALT, alanine aminotransferase; AST, aspartate aminotransferase; TG, triglyceride.

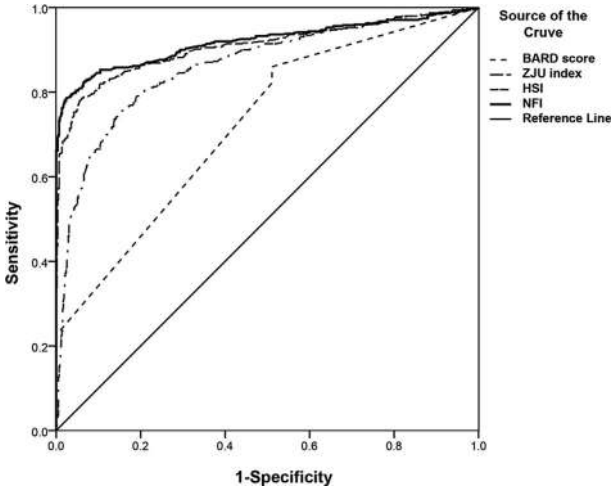


Fig. 2. AUROC of the BARD score, ZJU index, HSI, and NFI for prediction of non-alcoholic fatty liver disease. AUROC, area under the receiver-operating characteristic curve.

when the liver function or the imaging test was abnormal would people seek medical help. Although liver biopsy is regarded as the golden standard, it is accompanied by the risk of serious complications such as internal bleeding, bile leakage, and hematoma and infection, which limited its clinical application and was not the preferred method for NAFLD screening or diagnostic evaluation. Therefore, US, CT, or MRI-based methods are the main diagnostic tools for fatty liver.[38] Shear wave elastography (SWE) is an ultrasound-based technique, which has been applied to characterize liver fibrosis.[39] A recent meta-analysis demonstrated that SWE and MRI had a high diagnostic accuracy for staging fibrosis in NAFLD patients, with the summary sensitivity and specificity of SWE and MRI for detecting advanced fibrosis (AF) being 0.90 and 0.93, and 0.84 and 0.90, respectively, and the summary AUROC values using SWE and MRI for diagnosing AF were 0.95 and 0.96, respectively.[40] The results were consistent with those reported by Hamaguchi et al.,[41] who established a screening model of NAFLD with ultrasound imaging. Nevertheless, the drawbacks of imaging diagnosis are the high cost and inconvenience, so it is not suitable for large-scale population screening.

The FLI incorporates the results of biochemical tests and anthropometric data to predict steatosis.[28] It has been widely used to predict early-stage fatty liver and has been validated in many cohorts from different ethnic populations and countries.[30,31,42] However, WHO has proposed different cut-off values of BMI for the Asian populations from those for Western populations.[43] Yang et al.[25] reported that, when they used lower cut-off values of FLI than those used by Bedogni et al.,[28] the FLI could accurately identify ultrasonographic fatty liver in a large-scale population in Taiwan. Similar to the FLI, the BARD score[33] and HSI[32] are both simple and non-invasive models to predict fatty liver in Western populations, and they may not be suitable for the Chinese.

Table 4
Predictive values for NFI in the study population.

Variables	Low cut-off point (<31.0)	Intermediate (31.0–36.0)	High cut-off point (>36.0)
Total, *n* (%)	168 (14.7%)	561 (49%)	416 (36.3%)
NAFLD, *n* (%)	17 (10.1%)	123 (21.9%)	410 (98.6%)
Sensitivity	96.9% (95.0–98.1)		75.6% (72.1–79.2)
Specificity	25.4% (22.0–29.1)		98.9% (97.7–99.6)
LR_+	1.299 (1.236–1.364)		75.230 (33.890–167.00)
LR_-	0.122 (0.076–0.196)		0.243 (0.211–0.282)
PPV	–		98.6% (96.7–99.4)
NPV	89.9% (84.0–93.8)		–
Interpretation	Absence of NAFLD		Presence of NAFLD

Abbreviations: LR_+, positive likelihood ratio; LR_-, negative likelihood ratio; PPV, positive predictive value; NPV, negative predictive value.

The AUROC of the NFI was 0.919 (95% CI = 0.901–0.937), which was higher than that of the BARD score, HIS, and FLI in our participants. Furthermore, we found that the NFI had a better diagnostic accuracy than the ZJU index, and the AUROC of the ZJU index was 0.908 (95% CI = 0.889–0.926). When we analyzed the detection rate in participants of different age groups stratified by gender, the detection rate of NAFLD in males was nearly twice that of females below 50 years old, but the gap narrowed with the increase of age. Over 60 years old, the detection rate of NAFLD in females was nearly the same as in males, and similar results have been shown in another study.[44]

In this study, three components of MetS (BMI, FBG, and TG) were in the predictive model, suggesting that there is a close relationship between MetS and NAFLD. A growing body of evidence has demonstrated that the above-mentioned components of MetS are risk factors of NAFLD, and the larger the number of the above-mentioned components a patient has, the greater the effect on NAFLD.[45,46] A possible mechanism may be that IR plays a key role in the occurrence of these two diseases.[47] In individuals with MetS, IR may cause the storage and decomposition of lipids in the liver and promote fatty acid transportation from adipose tissues into the liver, leading to the synthesis of TG and secretion of very low-density lipoprotein in the liver, which result in fatty liver.[48]

Our study was conducted to detect NAFLD in a large Chinese population with standardized demographic, anthropometric, and laboratory measures. Our findings suggest that the NFI is accurate and can be used for the effective prediction of the presence of NAFLD. However, our study has some limitations. First, the study was a cross-sectional study and we did not match the participants for age or gender, so potential bias may exist and affect the results. Second, all the patients with NAFLD were diagnosed by US instead of liver biopsy. Third, factors such as physical activity and medical history were not included in this study, which may affect the BMI, FPG, and TG level in our study and result in bias in the final equation.

Table 3
A comparison of the area under the receiver-operating characteristic curves (AUROC) of each test for diagnosing non-alcoholic fatty liver disease (NAFLD).

Test	AUROC (95% CI)	Cut-off point	Sensitivity (%)	Specificity (%)	LR_+	LR_-
BARD score	0.722 (0.693–0.751)	0.500	86.0	48.9	1.683	0.286
ZJU index	0.908 (0.889–0.926)	34.775	78.0	94.5	14.182	0.233
HSI	0.862 (0.840–0.884)	32.945	75.3	85.2	5.088	0.290
NFI	0.919 (0.901–0.937)	35.465	78.5	97.6	32.708	0.220

Abbreviations: AUROC, area under the receiver-operating characteristic curve; LR_+, positive likelihood ratio; LR_-, negative likelihood ratio.
The BARD score was a weighted sum (BMI ≥28 = 1 point, AST/ALT >0.8 = 2 points, diabetes = 1 point).
The ZJU index was calculated as follows: BMI (kg/m^2) + FPG (mmol/L) + TG (mmol/L) + 3 × ALT (IU/L)/AST (IU/L) ratio (+2, female).
HSI was calculated as follows: 8 × ALT (IU/L)/AST (IU/L) + BMI (kg/m^2) (+2, if DM; +2, female).

Conclusion

In conclusion, we created a cost-effective NAFLD-screening model, named NFI, which showed a high accuracy for predicting NAFLD at early stages in a Chinese population. Moreover, it was easy to employ as BMI, FBG, TG, ALT, and AST were routine measurements in clinical practice. In our study, an NFI of <31.0 was ruled out and an NFI of >36.0 ruled in NAFLD with a higher diagnostic value. We hope that this could be directed toward high-risk groups identified by this model to help them change unhealthy dietary habits and lifestyles.

Author statements

Acknowledgments

The authors thank the study participants, collaborators, and staff.

Ethical approval

None sought.

Funding

This work was supported by the Ningbo Science and Technology Innovation Team Program under grant number 2014B82002, the National Natural Science Foundation of China under grant numbers 81471038 and 11504372, the Zhejiang Provincial Natural Science Foundation of China under grant number Y17C060002, and Fang Runhua Fund of Hong Kong, K. C. Wong Magna Fund in Ningbo University.

Competing interests

All authors have no conflict of interest to declare.

References

1. Benedict M, Zhang X. Non-alcoholic fatty liver disease: an expanded review. *World J Hepatol* 2017;**9**:715–32.
2. Stefan N, Haring HU, Cusi K. Non-alcoholic fatty liver disease: causes, diagnosis, cardiometabolic consequences, and treatment strategies. *Lancet Diab Endocrinol* 2019;**7**:313–24.
3. Jiang W, Huang S, Teng H, Wang P, Wu M, Zhou X, et al. Diagnostic accuracy of point shear wave elastography and transient elastography for staging hepatic fibrosis in patients with non-alcoholic fatty liver disease: a meta-analysis. *BMJ open* 2018;**8**:e021787.
4. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice guideline by the American Gastroenterological Association, American Association for the Study of Liver Diseases, and American College of Gastroenterology. *Gastroenterology* 2012;**142**:1592–609.
5. Adams LA, Lymp JF, St Sauver J, Sanderson SO, Lindor KD, Feldstein A, et al. The natural history of nonalcoholic fatty liver disease: a population-based cohort study. *Gastroenterology* 2005;**129**:113–21.
6. Williams CD, Stengel J, Asike MI, Torres DM, Shaw J, Contreras M, et al. Prevalence of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis among a largely middle-aged population utilizing ultrasound and liver biopsy: a prospective study. *Gastroenterology* 2011;**140**:124–31.
7. Pearson-Stuttard J, Zhou B, Kontis V, Bentham J, Gunter MJ, Ezzati M. Worldwide burden of cancer attributable to diabetes and high body-mass index: a comparative risk assessment. *Lancet Diab Endocrinol* 2018;**6**:e6–15.
8. Eslamparast T, Tandon P, Raman M. Dietary composition independent of weight loss in the management of non-alcoholic fatty liver disease. *Nutrients* 2017;**9**.
9. Ekstedt M, Hagstrom H, Nasr P, Fredrikson M, Stal P, Kechagias S, et al. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. *Hepatology* 2015;**61**:1547–54.
10. Aiello AM, Marques de Mello L, Souza Nunes M, Soares da Silva A, Nunes A. Prevalence of obesity in children and adolescents in Brazil: a meta-analysis of cross-sectional studies. *Curr Pediatr Rev* 2015;**11**:36–42.
11. Skinner AC, Ravanbakht SN, Skelton JA, Perrin EM, Armstrong SC. *Prevalence of obesity and severe obesity in US children, 1999-2016*. Pediatrics; 2018.
12. Zhang X, Zhang F, Yang J, Yang W, Liu W, Gao L, et al. Prevalence of overweight and obesity among primary school-aged children in Jiangsu Province, China, 2014-2017. *PLoS One* 2018;**13**:e0202681.
13. Ji CY, Cheng TO. Prevalence and geographic distribution of childhood obesity in China in 2005. *Int J Cardiol* 2008;**131**:1–8.
14. Song Y, Wang HJ, Ma J, Wang Z. Secular trends of obesity prevalence in urban Chinese children from 1985 to 2010: gender disparity. *PLoS One* 2013;**8**:e53069.
15. Anderson EL, Howe LD, Jones HE, Higgins JP, Lawlor DA, Fraser A. The prevalence of non-alcoholic fatty liver disease in children and adolescents: a systematic review and meta-analysis. *PLoS One* 2015;**10**:e0140908.
16. Yasui K, Hashimoto E, Komorizono Y, Koike K, Arii S, Imai Y, et al. Characteristics of patients with nonalcoholic steatohepatitis who develop hepatocellular carcinoma. *Clin Gastroenterol Hepatol: Off Clin Pract J Am Gastroenterol Assoc* 2011;**9**:428–33. quiz e50.
17. Rinella ME. Nonalcoholic fatty liver disease: a systematic review. *JAMA* 2015;**313**:2263–73.
18. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology* 2016;**64**:73–84.
19. Li Z, Xue J, Chen P, Chen L, Yan S, Liu L. Prevalence of nonalcoholic fatty liver disease in mainland of China: a meta-analysis of published studies. *J Gastroenterol Hepatol* 2014;**29**:42–51.
20. Fan JG, Farrell GC. Epidemiology of non-alcoholic fatty liver disease in China. *J Hepatol* 2009;**50**:204–10.
21. Targher G, Bertolini L, Rodella S, Zoppini G, Lippi G, Day C, et al. Non-alcoholic fatty liver disease is independently associated with an increased prevalence of chronic kidney disease and proliferative/laser-treated retinopathy in type 2 diabetic patients. *Diabetologia* 2008;**51**:444–50.
22. Targher G, Day CP, Bonora E. Risk of cardiovascular disease in patients with nonalcoholic fatty liver disease. *N Engl J Med* 2010;**363**:1341–50.
23. Targher G, Byrne CD. Non-alcoholic fatty liver disease: an emerging driving force in chronic kidney disease. *Nat Rev Nephrol* 2017;**13**:297–310.
24. Perumpail BJ, Khan MA, Yoo ER, Cholankeril G, Kim D, Ahmed A. Clinical epidemiology and disease burden of nonalcoholic fatty liver disease. *World J Gastroenterol* 2017;**23**:8263–76.
25. Yang BL, Wu WC, Fang KC, Wang YC, Huo TI, Huang YH, et al. External validation of fatty liver index for identifying ultrasonographic fatty liver in a large-scale cross-sectional study in Taiwan. *PLoS One* 2015;**10**:e0120443.
26. Pais R, Charlotte F, Fedchuk L, Bedossa P, Lebray P, Poynard T, et al. A systematic review of follow-up biopsies reveals disease progression in patients with non-alcoholic fatty liver. *J Hepatol* 2013;**59**:550–6.
27. Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, Bruckert E, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. *Gastroenterology* 2005;**128**:1898–906.
28. Bedogni G, Bellentani S, Miglioli L, Masutti F, Passalacqua M, Castiglione A, et al. The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population. *BMC Gastroenterol* 2006;**6**:33.
29. Huang X, Xu M, Chen Y, Peng K, Huang Y, Wang P, et al. Validation of the fatty liver index for nonalcoholic fatty liver disease in middle-aged and elderly Chinese. *Medicine* 2015;**94**:e1682.
30. Shen YN, Yu MX, Gao Q, Li YY, Huang JJ, Sun CM, et al. External validation of non-invasive prediction models for identifying ultrasonography-diagnosed fatty liver disease in a Chinese population. *Medicine* 2017;**96**:e7610.
31. Kim JH, Kwon SY, Lee SW, Lee CH. Validation of fatty liver index and lipid accumulation product for predicting fatty liver in Korean population. *Liver Int–Off J Int Assoc Study Liver* 2011;**31**:1600–1.
32. Lee JH, Kim D, Kim HJ, Lee CH, Yang JI, Kim W, et al. Hepatic steatosis index: a simple screening tool reflecting nonalcoholic fatty liver disease. *Dig Liver Dis–Off J Ital Soc Gastroenterol Ital Assoc Study Liver* 2010;**42**:503–8.
33. Harrison SA, Oliver D, Arnold HL, Gogia S, Neuschwander-Tetri BA. Development and validation of a simple NAFLD clinical scoring system for identifying patients without advanced disease. *Gut* 2008;**57**:1441–7.
34. Wang J, Xu C, Xun Y, Lu Z, Shi J, Yu C, et al. ZJU index: a novel model for predicting nonalcoholic fatty liver disease in a Chinese population. *Sci Rep* 2015;**5**:16494.
35. Levy JC, Matthews DR, Hermans MP. Correct homeostasis model assessment (HOMA) evaluation uses the computer program. *Diabetes Care* 1998;**21**:2191–2.
36. Dasarathy S, Dasarathy J, Khiyami A, Joseph R, Lopez R, McCullough AJ. Validity of real time ultrasound in the diagnosis of hepatic steatosis: a prospective study. *J Hepatol* 2009;**51**:1061–7.
37. Palmentieri B, de Sio I, La Mura V, Masarone M, Vecchione R, Bruno S, et al. The role of bright liver echo pattern on ultrasound B-mode examination in the diagnosis of liver steatosis. *Dig Liver Dis – Off J Ital Soc Gastroenterol Ital Assoc Study Liver* 2006;**38**:485–9.
38. Khov N, Sharma A, Riley TR. Bedside ultrasound in the diagnosis of nonalcoholic fatty liver disease. *World J Gastroenterol* 2014;**20**:6821–5.
39. Zheng J, Guo H, Zeng J, Huang Z, Zheng B, Ren J, et al. Two-dimensional shear-wave elastography and conventional US: the optimal evaluation of liver fibrosis and cirrhosis. *Radiology* 2015;**275**:290–300.
40. Xiao G, Zhu S, Xiao X, Yan L, Yang J, Wu G. Comparison of laboratory tests, ultrasound, or magnetic resonance elastography to detect fibrosis in patients with nonalcoholic fatty liver disease: a meta-analysis. *Hepatology* 2017;**66**:1486–501.

41. Hamaguchi M, Kojima T, Itoh Y, Harano Y, Fujii K, Nakajima T, et al. The severity of ultrasonographic findings in nonalcoholic fatty liver disease reflects the metabolic syndrome and visceral fat accumulation. *Am J Gastroenterol* 2007;**102**:2708–15.
42. Koehler EM, Schouten JN, Hansen BE, Hofman A, Stricker BH, Janssen HL. External validation of the fatty liver index for identifying nonalcoholic fatty liver disease in a population-based study. *Clin Gastroenterol Hepatol – Off J Pract J Am Gastroenterol Assoc* 2013;**11**:1201–4.
43. Consultation WHOE. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet* 2004;**363**: 157–63.
44. Farrell GC, Wong VW, Chitturi S. NAFLD in Asia–as common and important as in the West. *Nat Rev Gastroenterol Hepatol* 2013;**10**:307–18.
45. Zhang T, Zhang C, Zhang Y, Tang F, Li H, Zhang Q, et al. Metabolic syndrome and its components as predictors of nonalcoholic fatty liver disease in a northern urban Han Chinese population: a prospective cohort study. *Atherosclerosis* 2015;**240**:144–8.
46. Hamaguchi M, Takeda N, Kojima T, Ohbora A, Kato T, Sarui H, et al. Identification of individuals with non-alcoholic fatty liver disease by the diagnostic criteria for the metabolic syndrome. *World J Gastroenterol* 2012;**18**: 1508–16.
47. Asrih M, Jornayvaz FR. Metabolic syndrome and nonalcoholic fatty liver disease: is insulin resistance the link? *Mol Cell Endocrinol* 2015;**418**(Pt 1):55–65.
48. Puri P, Baillie RA, Wiest MM, Mirshahi F, Choudhury J, Cheung O, et al. A lipidomic analysis of nonalcoholic fatty liver disease. *Hepatology* 2007;**46**: 1081–90.

Public Health 202 (2022) 35–42

Contents lists available at ScienceDirect

Public Health

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

Original Research

Perceived changes in lifestyle behaviours and in mental health and wellbeing of elementary school children during the first COVID-19 lockdown in Canada

Katerina Maximova [a, b, c], Mohammad K.A. Khan [b], Julia Dabravolskaj [c], Laena Maunula [c], Arto Ohinmaa [c], Paul J. Veugelers [c, *]

[a] MAP Centre for Urban Health Solutions, Li Ka Shing Knowledge Institute, St. Michael's Hospital, 209 Victoria St, M5B 1T8, Toronto, Ontario, Canada
[b] Dalla Lana School of Public Health, University of Toronto, 155 College St Room 500, M5T 3M7, Toronto, Ontario, Canada
[c] School of Public Health, University of Alberta, 3-50E University Terrace, 8303 112 Street NW, T6G 1K4, Edmonton, Alberta, Canada

ARTICLE INFO

Article history:
Received 10 May 2021
Received in revised form
21 September 2021
Accepted 9 October 2021
Available online 20 October 2021

Keywords:
COVID-19
Children
Lifestyle behaviours
Mental health and wellbeing
School health
Public health

ABSTRACT

Objectives: The closure of schools to prevent the spread of COVID-19 prompted concerns of deteriorating lifestyle behaviours, mental health, and wellbeing of children, particularly those in socioeconomically disadvantaged settings. We assessed changes in lifestyle behaviours (physical activity, screen time, eating habits and bed/wake-up times), mental health and wellbeing during the first lockdown in Spring 2020 as perceived by school children from disadvantaged settings, and examined determinants of these changes.
Study design: Cross-sectional study.
Methods: We surveyed 1095 grade 4 to 6 students (age 9–12 years) from 20 schools in socioeconomically disadvantaged communities in northern Canada. Students reported on changes in lifestyle behaviours, mental health and wellbeing during the lockdown. Determinants of these perceived changes were examined in multivariable regression models.
Results: A majority of students reported declines in physical activity, having late bed/wake-up times, and modest improvements in mental health and wellbeing. Many students reported increases rather than decreases in screen time and snacking. Positive attitudes toward being active, eating healthy, going to sleep on time and being healthy were strongly associated with maintaining healthy lifestyle behaviours during the lockdown. Positive attitudes toward active and healthy living and healthy lifestyle behaviours were associated with maintaining positive mental health and wellbeing during the lockdown.
Conclusions: The considerable changes in lifestyle behaviors, superimposed on the pre-existing burden of unhealthy lifestyle behaviours, put this generation of children at increased risk for future chronic disease. Findings call for effective health promotion of active and healthy lifestyles to benefit both physical and mental health.

Introduction

The COVID-19 pandemic, declared by the World Health Organization on March 11, 2020,[1] resulted in the implementation of drastic public health measures that affected large populations of school-aged children.[2,3] While effective at reducing the viral spread, these measures prompted concerns regarding children's lifestyle behaviours (physical activity, screen time, eating habits and sleep patterns) and mental health and wellbeing.[4]

Healthy lifestyles, mental health and wellbeing are essential for physical, social and emotional development of children.[5] Unhealthy lifestyles and psychological problems at a young age may be difficult to reverse and will track into adulthood,[6–9] thereby predisposing children to a range of chronic diseases and mental illness later in life.[10] A recent survey of Canadian parents revealed that only 4.8% of children (5–11 years) and 0.6% of adolescents (12–17 years old) adhered to physical activity and sedentary behaviour recommendations during the early pandemic period, although 71.1% were meeting sleep recommendations.[11] In addition, some

* Corresponding author. School of Public Health, University of Alberta, 3-50E University Terrace, 8303 112 Street NW, Edmonton, Alberta, T6G 1K4 Canada. Tel.: +(780) 492 9095.
E-mail address: paul.veugelers@ualberta.ca (P.J. Veugelers).

https://doi.org/10.1016/j.puhe.2021.10.007

emerging studies suggest, albeit not consistently, that mental health declined among children during the first COVID-19 lockdown.[12–14]

The health and wellbeing of socioeconomically disadvantaged children during the pandemic is of particular concern.[15] The prevalence of poor lifestyles and mental illness is already high among these children, and studies report on low adherence to lifestyle recommendations during the lockdown in socioeconomically disadvantaged settings.[16] In addition, the pandemic-caused stressors (household food insecurity, parental job loss,[17] disruptive family dynamics[18]) may disproportionally affect the physical and mental health of vulnerable children. To date, existing evidence has come predominantly from online surveys in convenience samples of parents rather than children, and children's perspectives on their lifestyle behaviours, mental health and wellbeing have not yet been heard. Among elementary school children residing in socioeconomically disadvantages communities, we (1) examined the perceived changes in lifestyle behaviours, mental health and wellbeing and (2) assessed the role of attitudes toward active and healthy living and other determinants of these changes.

Methods

During in-person learning in November–December 2020 and January–February 2021, we invited 1340 students in grades 4 to 6 from 20 schools located in British Columbia, Alberta, Manitoba, and Northwest Territories to participate in the survey. All 20 schools are part of the Alberta Project Promoting healthy Living for Everyone (APPLE) in Schools project – an innovative, internationally recognized, not-for-profit health promotion program targeting children from socioeconomically disadvantaged communities. Grounded in a Comprehensive School Health approach, the program promotes healthy lifestyle behaviours, mental health and wellbeing by transforming the school's culture to "make the healthy choice the easy choice".[19,20]

Data were collected in school during regular class time, with research assistants prompting the survey questions projected on the whiteboard through Zoom. A total of 1095 students completed the survey, with the participation rate of 81.6%. The Health Research Ethics Board of the University of Alberta (Pro00061528) and participating school boards approved all the procedures.

As part of the pandemic response in Canada, in-school learning in all participating schools (among others) was suspended mid-March 2020. Although school buildings were closed, online learning continued for the duration of the 2019/2020 school year. Grade K-6 students were permitted to attend in-person classes when the participating schools re-opened in September 2020 with enhanced public health measures in place (e.g. cohorting, masking, physical distancing)[21–23] and an early and extended Christmas break.

Students were asked to compare their activity levels during the Spring 2020 lockdown with their activity levels before this lockdown with respect to 24 common physical activities derived from the PAQ-C.[24] Response categories were adapted to reflect changes in activity levels, as perceived by the students, during the COVID-19 lockdown. Response categories (Less than/About the same/More than before schools got closed) were assigned a score of '-1', '0' and '1', averaged and dichotomized using '0' as the cut-off value, with lower values indicating less physical activity during the lockdown, and '0' and above more physical activity. The same response categories were provided for the questions about perceived changes in time spent playing video games and using a cellphone, and in the number of meals and snacks. With respect to sleep, students reported their wake-up time and bedtime on weekdays. In accordance with the Canadian 24-h Movement Guidelines for children and youth,[25] we coded responses for weekday wake-up time after 10:00 am or bedtime after 11:00 pm as having late bed/wake-up times. The surveys included 11 questions related to mental health and wellbeing (details described in our previous work[26]), with response categories adapted to capture perceived changes during the lockdown (More/About the same/Less than before schools got closed). The response categories were assigned a score of '-1', '0' and '1' for positively stated items and reverse coded for negatively stated items. The cumulative score was created and dichotomized using '0' as the cut-off value, with values above '0' representing better mental health and wellbeing, and between '-12' and '0' (inclusive) worse mental health and wellbeing during the lockdown.

Attitudes toward active and healthy living were assessed by a series of questions: "How much do you care about being healthy? Being physically active? Eating healthy foods? Going to sleep on time?" The response categories were 'very much', 'quite a lot', 'a little bit', and 'not at all.' We considered these attitudes as potential determinants of changes in lifestyle behaviours, mental health and wellbeing based on their utility to help prioritize future prevention intervention efforts. Given existing differences in lifestyle behaviours by gender, grade,[27] race/ethnicity, and socioeconomical status,[28] we considered the student's gender (girl, boy), grade level (4, 5, 6), languages spoken (English only, English and Indigenous language(s), and English and other language(s)), region of residence (rural, small population centre, medium, and large population centre[29]) as potential confounders. In addition, we adjusted for social and material deprivation indices that were derived from Canada Census 2016 data based on postal codes of APPLE Schools included in the analysis (detailed description of and procedures for calculating these indices can be found elsewhere[30]). Higher quintiles for both indices indicate more deprived areas.[31,32] Finally, we considered the length of time since school reopening in September 2020 (<3 months, ≥3 months) as a confounder to account for the possibility that student's lifestyle behaviours and mental health and wellbeing three or more months after reopening could have returned to the pre-pandemic levels.

Statistical analyses

First, multivariable logistic regression models were used to examine the association of the attitudes toward being healthy with perceived changes in physical activity, playing video games, cellphone use, maintaining and adopting good meal and snack routines, not having late bed/wake-up times, and maintaining positive mental health and wellbeing while adjusting for potential confounders. Gender-stratified exploratory factor analysis with varimax (orthogonal) rotation was employed to extract latent factors that maximized the explained variance. After examination of the scree plot and based on the Kaiser criterion (i.e. eigenvalue >1), three clusters were identified separately for girls and boys to reflect gender differences in clustering of the responses (Table S1). Factor scores were then calculated for each of the clusters using the Bartlett's test of sphericity,[33] dichotomized and considered as outcomes in the multivariable logistic regression models. Before analyses, missing values were imputed for lifestyle behaviors and mental health and wellbeing based on 'multivariate imputation by chain equation (MICE)'.[34] Approximately 80% of students provided responses to each of the 20 questions (11 questions on mental health and wellbeing, five on screen time, two on eating habits, and two on sleep), and 90% completed at least 19 of the items. Data from 31 students who did not respond to 10 or more of the 20 questions were excluded from analyses. Fixed effects regression modeling was applied instead of mixed effects models since the intra-class correlation was below 0.02 for all models.[35] In addition, we

considered changes in lifestyle behaviours as potential correlates of changes in mental health and wellbeing.

Results

Table 1 shows the characteristics of the 1095 participants and their schools. There were more girls (n = 557) than boys (n = 538) and more grade 6 (n = 400) relative to grade 4 (n = 312) and 5 (n = 383) students participating. Twenty percent of students reported speaking one or more Indigenous languages and 11% speaking another language in addition to English. In terms of students' attitudes toward active and healthy living, the majority of girls and boys cared 'very much' or 'quite a lot' about being healthy, physically active and eating healthy, although less than half reported they cared about good sleep.

About two-thirds of girls and boys (62% and 64%, respectively) recalled their physical activity levels to be lower during the lockdown than before the lockdown (Fig. 1). Almost two-thirds of boys (64%) reported they spent more time playing video games during the lockdown, whereas about one-quarter (24%) reported this to be about the same and 12% spent less time playing video games during the lockdown. Almost half the girls (46%) and more than a third of boys (38%) reported using a cellphone more, 23% and 27% of girls and boys reported using it less than before, and the rest of the students reported no change. Almost half of girls (48%) and 37% of boys reported snacking more than 14% and 16% of girls and boys who reported snacking less during the lockdown. Sixty-eight percent of girls and 67% of boys reported late bed/wake-up times (Fig. 1).

While 44% and 31% of girls and boys perceived, on average, their mental health and wellbeing to be worse, the majority perceived their mental health and wellbeing to be better during the lockdown (Fig. 1). However, responses to the individual questions were mixed (Fig. 2). For example, 35% of girls and 27% of boys indicated they felt lonely more often compared to 33% and 41% of girls and boys who felt lonely less often during the lockdown than before the schools got closed. Thirty-two percent of students reported feeling lonely the same as before the lockdown. Almost half of girls (48%) and 36% of boys indicated they were more bored than 22% and 33% of girls and boys reporting they were less bored during the lockdown than before the schools got closed. At the same time, most of girls (47%) and boys (58%) reported enjoying their time at home more than before the lockdown, while the rest enjoyed their time at home less and about the same (Fig. 2).

Students who cared about being physically active were more likely to report no increases in time playing video games and using a cellphone and were more likely to maintain positive mental health (Table 2). Similarly, those who cared about eating

Table 1
Characteristics of students and schools that participated in the study, Canada, 2020/21.

Student Characteristics	Girls (n = 557)	Boys (n = 538)	Total (n = 1095)
	n (%)	n (%)	n (%)
Grade Level			
Grade 4	161 (29)	151 (28)	312 (28)
Grade 5	180 (32)	203 (38)	383 (35)
Grade 6	216 (39)	184 (34)	400 (37)
Attitudes toward being healthy			
Care[a] about being healthy	486 (87)	460 (85)	926 (86)
Care[a] about being physically active	429 (77)	409 (76)	838 (77)
Care[a] about eating healthy	416 (75)	376 (69)	792 (72)
Care[a] about good sleep	268 (49)	235 (44)	503 (46)
Language(s) spoken			
English only	387 (69)	374 (70)	761 (69)
English and Indigenous	109 (20)	110 (20)	219 (20)
English and other	61 (11)	54 (10)	115 (11)
School characteristics	Schools (n = 20)		Students (n = 1095)
	n (%)		n (%)
Time since reopening			
≤3 months	13 (65)		672 (61)
>3 months	7 (35)		423 (39)
Region of residence[b]			
Rural	8 (40)		165 (15)
Small PC	7 (35)		583 (53)
Medium PC	2 (10)		188 (17)
Large PC	3 (15)		159 (15)
Material deprivation quintile			
1 (least deprived)	2 (10)		66 (6)
2	4 (20)		222 (20)
3	5 (25)		447 (41)
4	5 (25)		291 (27)
5 (most deprived)	4 (20)		69 (6)
Social deprivation quintile			
1 (least deprived)	5 (25)		224 (20)
2	6 (30)		402 (37)
3	4 (20)		197 (18)
4	3 (15)		164 (15)
5 (most deprived)	2 (10)		108 (10)

PC: population centre.
[a] Percent of students who responded 'very much' or 'quite a lot'.
[b] Rural refers to a community with <1000 population, small PC to 1000–29,999 population, medium PC to 30,000–99,999, large PC refers to population of 100,000 or more.[29]

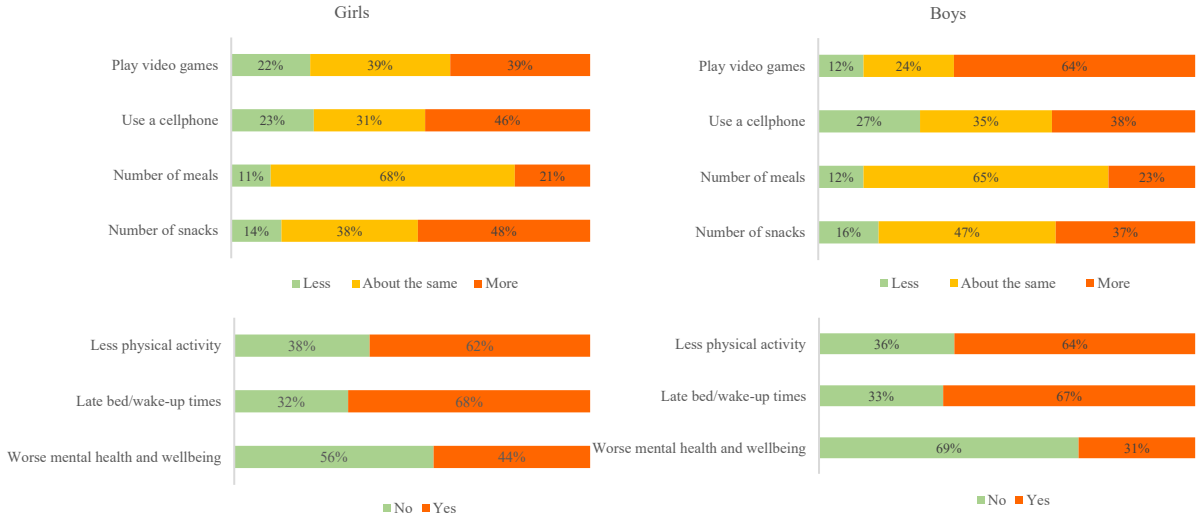


Fig. 1. Changes in physical activity, time playing video games, time using the cellphones, meal and snack frequency, bed/wake-up times, and mental health and wellbeing during vs before the lockdown.

healthy were more likely to report no increases in snacking during the lockdown and were likely to consume same amount or fewer snacks, while those who cared about their sleep and being healthy were less likely to report late bed/wake-up times during the lockdown. Gender-stratified associations are presented in Tables S2 and S3.

Tables 3 and 4 report on the associations of changes in lifestyle behaviours with changes in mental health and wellbeing in girls and boys, respectively. Girls who were more physically active during than before the lockdown were less likely to experience 'internalizing and functioning problems, tiredness and loneliness' and more likely to have a 'positive outlook on future and time during lockdown' relative to those who were less physically active (Table 3). In girls, spending less or the same amount of time playing video games was associated with a higher likelihood of maintaining positive mental health and wellbeing during the lockdown and being bored and lonely. In addition, having more meals was associated with having a 'positive outlook on future and time during the lockdown'. Similar to girls, boys who were physically active during the lockdown were more likely to have a 'positive outlook on future and time during the lockdown'. Boys who reported having more meals and fewer snacks were more likely to experience positive mental health and wellbeing during the lockdown (Table 4).

Discussion

The present study reported on changes in the lifestyle behaviours and mental health and wellbeing, as perceived by elementary school–aged children living in socioeconomically disadvantaged communities, during the Spring 2020 COVID-19 lockdown.

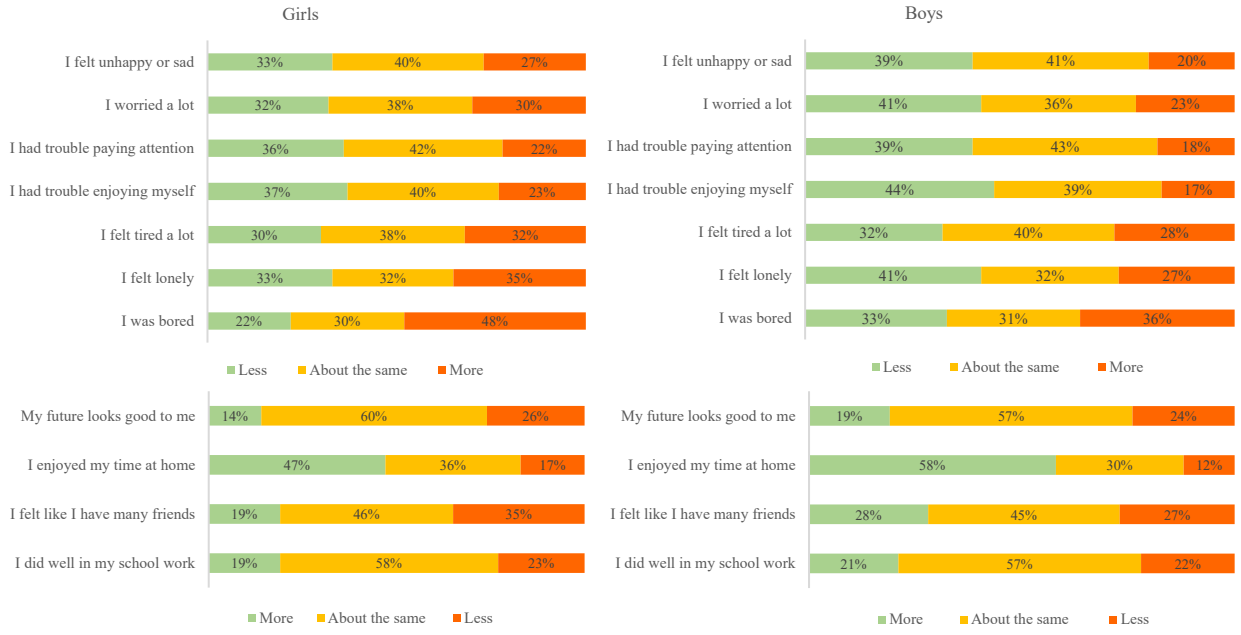


Fig. 2. Changes in mental health and wellbeing during vs before the lockdown.

Table 2
Associations[a] of attitudes toward active and healthy living with changes in physical activity, sedentary behaviours, healthy eating, sleep, mental health and wellbeing during vs before the lockdown, Canada, 2020/21.

	More physical activity (vs less)	Same/less time playing video games (vs more)	Same/less time on cell phone (vs more)	Same/more meals (vs less)	Same/fewer snacks (vs more)	No late bed/wake-up times (vs late)	Better mental health and wellbeing (vs worse)
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Care about being healthy							
A little/not at all	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Very much/quite a lot	1.12 (0.72, 1.78)	1.20 (0.77, 1.87)	0.94 (0.61, 1.42)	1.39 (0.70, 2.65)	0.83 (0.53, 1.28)	1.69 (1.13, 2.51)	1.13 (0.71, 1.79)
Care about being active							
A little/not at all	Ref	Ref	Ref				Ref
Very much/quite a lot	1.18 (0.83, 1.68)	2.12 (1.5, 3.00)	1.86 (1.35, 2.59)				1.39 (1.06, 1.82)
Care about eating healthy							
A little/not at all				Ref	Ref		Ref
Very much/quite a lot				0.67 (0.39, 1.12)	1.44 (1.05, 1.98)		1.29 (0.91, 1.81)
Care about good sleep							
A little/not at all						Ref	Ref
Very much/quite a lot						2.15 (1.63, 2.84)	0.85 (0.60, 1.19)

OR: odds ratio; 95% CI: 95% confidence interval; Ref: reference category.
[a] All estimates are adjusted for the student characteristics (gender, grade level, language(s) spoken) and school/community characteristics (social deprivation, material deprivation, region of residence, and time since reopening of schools).

Children reported to be less physically active, to spend more time playing video games (boys) and using cellphones (girls), and to snack more. The majority of students reported having late bed or wake-up times on weekdays during the lockdown. In contrast, the majority of students found their mental health and wellbeing to be better during the lockdown. Moreover, positive attitudes toward being active, eating healthy, going to sleep on time and being healthy were strongly associated with maintaining healthy lifestyle behaviours during the lockdown. Last, positive attitudes were also associated with maintaining positive mental health and wellbeing during the lockdown.

Several studies reported on a considerable disengagement in physical activity during the lockdown.[36–40] While this appears true for most of our respondents, it is noteworthy that changes for the better were reported by more than one-third of girls and boys. In a study of lifestyle behaviours in Irish adolescents (12–18 years old) during the first lockdown,[41] 20% and 30% of students reported exercising more and about the same, respectively, than 50% exercising less than pre-pandemic. The majority of boys and girls reported an increase in time playing video games and using cellphones, respectively. The decrease in physical activity combined with an increase in screen time is concerning. However, in the context of the pandemic where physical distancing is paramount, and multiple and repeat lockdowns are enforced; sedentary activities that involve peer/social interaction might have favourable effects on emotional wellbeing because students might use video gaming to connect with peers.[42] Health promotion messages should seek a balance between encouraging active lifestyles while affording children the opportunities for social interaction through online mediums, particularly those that encourage physical activity (e.g. exergaming).

Participants reported snacking more during the lockdown, which is consistent with emerging literature.[43,44] Despite the substantial increase in the consumption of processed foods since

Table 3
Associations of changes in lifestyle behaviours with changes in mental health and wellbeing and its subgrouping during vs before the lockdown among girls, Canada, 2020/21.

	Better mental health and wellbeing (vs worse)	Internalizing and functioning problems, tired and lonely (vs not)	Positive outlook (vs not)	Bored and lonely (vs not)
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Physical activity				
Less than before	Ref	Ref	Ref	Ref
More than before	0.66 (0.45, 0.97)	0.52 (0.36, 0.76)	1.69 (1.13, 2.55)	0.81 (0.55, 1.18)
Video games				
Less than/same as before	1.61 (1.09, 2.37)	1.25 (0.85, 1.83)	1.16 (0.78, 1.72)	1.54 (1.04, 2.27)
More than before	Ref	Ref	Ref	Ref
Cellphone use				
Less than/same as before	1.18 (0.8, 1.73)	1.01 (0.69, 1.47)	1.47 (0.99, 2.18)	1.86 (1.27, 2.73)
More than before	Ref	Ref	Ref	Ref
Number of meals				
Less than/same as before	Ref	Ref	Ref	Ref
More than before	1.66 (0.93, 2.99)	1.02 (0.57, 1.81)	1.78 (1, 3.18)	1.42 (0.79, 2.58)
Number of snacks				
Less than/same as before	1.32 (0.9, 1.94)	1.3 (0.89, 1.89)	1.19 (0.81, 1.77)	1.06 (0.72, 1.54)
More than before	Ref	Ref	Ref	Ref
Late bed/wake-up times				
No	Ref	Ref	Ref	Ref
Yes	0.74 (0.5, 1.09)	0.77 (0.52, 1.14)	0.83 (0.56, 1.24)	0.67 (0.45, 0.99)

OR: odds ratio; 95% CI: 95% confidence interval; Ref: reference category.
All estimates are adjusted for the following student characteristics (grade level, language(s) spoken) and school/community characteristics (social deprivation, material deprivation, region of residence, and time since reopening of schools).

Table 4
Associations of changes in lifestyle behaviours with changes in mental health and wellbeing and its subgrouping during vs before the lockdown among boys, Canada, 2020/21.

	Better mental health and wellbeing (vs worse)	Bored, tired and lonely (vs not)	Internalizing and functioning problems (vs not)	Positive outlook (vs not)
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Physical activity				
Less than before	Ref	Ref	Ref	Ref
More than before	1.18 (0.77, 1.82)	1.05 (0.71, 1.54)	0.69 (0.47, 1.02)	2.06 (1.39, 3.06)
Video games				
Less than/same as before	0.82 (0.53, 1.27)	0.86 (0.57, 1.29)	1.04 (0.7, 1.56)	0.65 (0.43, 0.97)
More than before	Ref	Ref	Ref	Ref
Cellphone use				
Less than/same as before	1.2 (0.78, 1.85)	1.2 (0.81, 1.78)	1.08 (0.73, 1.59)	0.8 (0.54, 1.19)
More than before	Ref	Ref	Ref	Ref
Number of meals				
Less than/same as before	Ref	Ref	Ref	Ref
More than before	1.92 (1.08, 3.39)	1.38 (0.8, 2.39)	1.11 (0.64, 1.92)	0.9 (0.52, 1.57)
Number of snacks				
Less than/same as before	1.83 (1.2, 2.78)	1.31 (0.9, 1.92)	1.22 (0.83, 1.78)	1.08 (0.73, 1.59)
More than before	Ref	Ref	Ref	Ref
Late bed/wake-up times				
No	Ref	Ref	Ref	Ref
Yes	1.44 (0.93, 2.26)	1.16 (0.78, 1.73)	1.04 (0.71, 1.55)	1.11 (0.75, 1.66)

OR: odds ratio; 95% CI: 95% confidence interval; Ref: reference category.
All estimates are adjusted for the following student characteristics (grade level, language(s) spoken) and school/community characteristics (social deprivation, material deprivation, region of residence, and time since reopening of schools).

the start of the pandemic,[45] there are also reports that many families used the lockdown as an opportunity to steer their eating habits toward healthier options.[46] This underlines the importance of the health promotion initiatives focusing on healthy eating habits in the family setting.

Our findings of late bed/wake-up times during the lockdown are consistent with emerging literature,[47] but are not necessarily alarming. Most schools provided flexible school hours during the lockdown, and a shift toward later bedtime among children was shown to be accompanied by longer sleep duration, improved sleep quality and less daytime sleepiness during the lockdown.[48] Adopting a flexible school time schedule during the lockdown may help ensure children meet the recommended number of hours of sleep.[49]

Evidence on the impact of the COVID-19 pandemic on mental health and wellbeing in children is equivocal, and our study also reports mixed findings. While some children reported their mental health and wellbeing to worsen during the lockdown, most students reported positive changes. These results are seemingly not consistent with studies showing modest adverse impact of the pandemic on children and youth's mental health and wellbeing.[50,51] For example, in a survey of 166 grade 4 students in South Korea, Choi et al.[52] reported an increase in stress levels along with unchanged life satisfaction, underscoring the importance of high quality parent–child relationship in supporting children's mental health and wellbeing during the pandemic. Cultural and contextual factors, such as resiliency in small communities in Canada,[53,54] may also underlie the findings in the current study. All participating schools are part of the APPLE Schools program that takes a Comprehensive School Health approach to promoting healthy lifestyle behaviours and mental wellness. Pre-pandemic research had shown that this program was effective in increasing physical activity levels,[55–57] reducing screen time, improving vegetables and fruit consumption, and preventing excess body weight.[58] APPLE Schools continued the delivery of their programming when school buildings were closed, connecting directly with students and parents, sharing resource/activities promoting healthy lifestyle, distributing exercise equipment, providing healthy food hampers, offering online activities (e.g. guided meditations), support from mental health therapists, among others. These activities could have helped students to weather the adverse effects of the lockdown, particularly on mental health and wellbeing given the emphasis of the APPLE Schools programming on promotion of mental health and wellbeing. Moreover, APPLE Schools programming also targets the attitudes of students. In this study, we revealed that students with positive attitudes toward active and healthy living, were more likely to maintain healthy lifestyle behaviors and mental health and wellbeing.

Most of the emerging literature on lifestyle behaviours and mental health and wellbeing of children is based on parental reports, while our study gathered information directly from students. Grade 4 to 6 students have appropriate literacy level to complete surveys and may be better at responding to questions on changes in lifestyle behaviours and mental health and wellbeing than their parents.[59] Another strength of this study is a very high response rate achieved by providing flexibility to teachers in scheduling the survey administration and coordinating on the day when most of the students were available. However, there are several limitations, including the cross-sectional design of the study. Longitudinal studies are required to study long-term health impacts that might result from the negative lifestyle changes associated with the pandemic-related public health measures.[60] Particular attention should be paid to socioeconomically disadvantaged groups and communities because the pandemic has exacerbated pre-existing inequalities, with evidence emerging that a range of adversities (e.g. financial burden, access to basic necessities) are maintained over time.[61] In addition, the surveys were administered several months after the lockdown, which may have affected participants' recall of the changes in lifestyle behaviours and mental health and wellbeing.

Conclusion

This study in the general population of elementary school–aged children from socioeconomically disadvantaged communities demonstrated considerable changes in physical activity, screen

time, eating habits and bed/wake-up times, albeit modest changes in mental health and wellbeing. Considering multiple and repeat lockdowns and the negative changes in lifestyles observed in this study population despite the ongoing APPLE Schools programming, investments in health promotion are critical to avoid a cascade of negative health consequences in the decades ahead.

Author statements

Acknowledgments

The authors thank the students, parents/guardians, and school principals for their participation in the research. They further thank teachers, school health facilitators and champions, project assistants, the APPLE Schools staff for facilitating the research, and specifically Tina Skakun, Katherine Dekker and Landra Walker for their major roles in coordinating and conducting the data collection.

Ethical approval

The Health Research Ethics Board of the University of Alberta (Pro00061528) and participating school boards approved all the procedures.

Funding

The Public Health Agency of Canada (grant #1516-HQ-000071) along with other partners (for details see[20]) funded the APPLE Schools programming in the Northern communities. The present study was supported by operating funds from the Canadian Institutes for Health Research to KM and PJV (grant #172685). KM holds the Murphy Family Foundation Chair in Early Life Interventions.

Competing interests

The authors have no real or perceived competing interests to disclose.

Authors' contributions

All listed authors contributed to study design, drafted and revised the article, and gave their final approval of the version submitted for publication. KM and PJV conceptualized the study and methodology, and secured funding and resources. MKAK, KM, and PJV developed a statistical analysis plan, and MKAK conducted all data analyses. JD and PJV accessed and verified the data and wrote the original draft. All authors reviewed and approved the final manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2021.10.007.

References

1. World Health Organization. *WHO Director-General's opening remarks at the media briefing on COVID-19*. 11 March 2020. Available 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.
2. Public Health Agency of Canada. *Community-based measures to mitigate the spread of coronavirus disease (COVID-19) in Canada*. 2021. Available from: https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection/health-professionals/public-health-measures-mitigate-covid-19.html.
3. United Nations. *Policy brief: education during COVID-19 and beyond*. 2020 Aug. Available from: https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2020/08/sg_policy_brief_covid-19_and_education_august_2020.pdf.
4. Zemrani B, Gehri M, Masserey E, Knob C, Pellaton R. A hidden side of the COVID-19 pandemic in children: the double burden of undernutrition and overnutrition. *Int J Equity Health* 2021 Jan 22;**20**(1):44.
5. Macera CA. Promoting healthy eating and physical activity for a healthier nation. *Cent Dis Control Prev* 2005:1—22.
6. Biddle SJH, Pearson N, Ross GM, Braithwaite R. Tracking of sedentary behaviours of young people: a systematic review. *Prev Med* 2010 Nov 1;**51**(5):345—51.
7. Lake AA, Mathers JC, Rugg-Gunn AJ, Adamson AJ. Longitudinal change in food habits between adolescence (11—12 years) and adulthood (32—33 years): the ASH30 Study. *J Public Health* 2006 Mar 1;**28**(1):10—6.
8. Lien N, Lytle LA, Klepp K-I. Stability in consumption of fruit, vegetables, and sugary foods in a cohort from age 14 to age 21. *Prev Med* 2001 Sep 1;**33**(3):217—26.
9. Telama R. Tracking of physical activity from childhood to adulthood: a review. *Obes Facts* 2009;**2**(3):187—95.
10. Loewen OK, Maximova K, Ekwaru JP, Faught EL, Asbridge M, Ohinmaa A, et al. Lifestyle behavior and mental health in early adolescence. *Pediatrics* 2019;**143**(5):e20183307.
11. Moore SA, Faulkner G, Rhodes RE, Brussoni M, Chulak-Bozzer T, Ferguson LJ, et al. Impact of the COVID-19 virus outbreak on movement and play behaviours of Canadian children and youth: a national survey. *J Behav Nutr* 2020;**17**(1):85.
12. Bignardi G, Dalmaijer ES, Anwyl-Irvine AL, Smith TA, Siugzdaite R, Uh S, et al. Longitudinal increases in childhood depression symptoms during the COVID-19 lockdown. *Arch Dis Child* 2020 Dec 9:1—7. 0.
13. Cost KT, Crosbie J, Anagnostou E, Birken CS, Charach A, Monga S, et al. Mostly worse, occasionally better: impact of COVID-19 pandemic on the mental health of Canadian children and adolescents. *Eur Child Adolesc Psychiatr* 2021 Feb 26:1—14. https://doi.org/10.1007/s00787-021-01744-3.
14. Ravens-Sieberer U, Kaman A, Erhart M, Devine J, Schlack R, Otto C. Impact of the COVID-19 pandemic on quality of life and mental health in children and adolescents in Germany. *Eur Child Adolesc Psychiatr* 2021 Jan:1—11. https://doi:10.1007/s00787-021-01726-5.
15. Bambra C, Riordan R, Ford J, Matthews F. The COVID-19 pandemic and health inequalities. *J Epidemiol Community Health* 2020 Nov 1;**74**(11):964—8.
16. Guerrero MD, Vanderloo LM, Rhodes RE, Faulkner G, Moore SA, Tremblay MS. Canadian children's and youth's adherence to the 24-h movement guidelines during the COVID-19 pandemic: a decision tree analysis. *J Sport* 2020;**9**(4):313—21.
17. Statistica Canada. *Labour market impacts of COVID-19 on Indigenous people: March to August 2020*. 2020. Available from: https://www150.statcan.gc.ca/n1/pub/45-28-0001/2020001/article/00085-eng.htm.
18. Gassman-Pines A, Ananat EO, Fitz-Henley 2nd J. COVID-19 and parent-child psychological well-being. *Pediatrics* 2020;**146**(4). e2020007294.
19. Pan-Canadian Joint Consortium for School Health. *Comprehensive school health framework*. 2016. Available from: http://www.jcsh-cces.ca/images/What_is_Comprehensive_School_Health_-_2-pager_-_July_2016.pdf.
20. APPLE Schools. *APPLE Schools is an innovative school-focused health promotion initiative*. 2021. Available from: https://www.appleschools.ca/.
21. Alberta Government. *School re-entry plan*. 2020 Aug. Available from: https://open.alberta.ca/dataset/a3910dd5-d52f-4a7d-821f-a381002419e9/resource/03c07743-31bc-46d1-9c65-b39335c4f3ad/download/edu-2020-2021-school-re-entry-plan-2020-0827.pdf.
22. Manitoba Government. *Welcoming our students back: restoring safe schools*. 2020 Jun. Available from: https://www.gov.mb.ca/asset_library/en/covid/k-12-reopeningplan.pdf.
23. Government of Northwest Territories. Reopening NWT schools safely. Available from: https://www.gov.nt.ca/covid-19/sites/covid/files/resources/reopening_nwt_schools_safely_plan_for_2020-21_eng.pdf.
24. Kowalski KC, Crocker PRE, Donen RM. *The physical activity questionnaire for older children (PAQ-C) and adolescents (PAQ-A) manual. Saskatoon, Saskatchewan, Canada*. Saskatoon: College of Kinesiology, University of Saskatchewan; 2004. p. 1—37.
25. Tremblay MS, Carson V, Chaput J-P, Connor Gorber S, Dinh T, Duggan M, et al. Canadian 24-hour movement guidelines for children and youth: an integration of physical activity, sedentary behaviour, and sleep. *Appl Physiol Nutr Metabol* 2016 Jun;**41**(6):S311—27. Suppl. 3.
26. Maximova K, Khan MKA, Austin SB, Kirk SFL, Veugelers PJ. The role of underestimating body size for self-esteem and self-efficacy among grade five children in Canada. *Ann Epidemiol* 2015 Oct 1;**25**(10):753—9.
27. Public Health Agency of Canada. *Health behaviour in school-aged children in Canada: focus on relationships*. 2016 Mar. Available from: https://healthycanadians.gc.ca/publications/science-research-sciences-recherches/health-behaviour-children-canada-2015-comportements-sante-jeunes/index-eng.php#c6.
28. Armstrong S, Wong CA, Perrin E, Page S, Sibley L, Skinner A. Association of physical activity with income, race/ethnicity, and sex among adolescents and young adults in the United States: findings from the National Health and Nutrition Examination Survey, 2007-2016. *JAMA Pediatr* 2018 Aug 1;**172**(8):732—40.

29. Statistics Canada. *Population centre and rural area classification 2016*. 2017. Available from: https://www.statcan.gc.ca/eng/subjects/standard/pcrac/2016/introduction.
30. Institut National de Santé Publique du Québec. *Material and social deprivation index*. 2019. Available from: https://www.inspq.qc.ca/en/deprivation/material-and-social-deprivation-index.
31. Pampalon R, Raymond G. A deprivation index for health and welfare planning in Quebec. *Chron Dis Can* 2000;**21**(3):104–13.
32. Pampalon R, Hamel D, Gamache P, Simpson A, Philibert MD. Validation of a deprivation index for public health: a complex exercise illustrated by the Quebec index. *CDIC* 2014;**34**(1). Available from: https://www.canada.ca/en/public-health/services/reports-publications/health-promotion-chronic-disease-prevention-canada-research-policy-practice/vol-34-no-1-2014/validation-deprivation-public-health-complex-exercise-illustrated-quebec.html.
33. Bartholomew DJ, Deary IJ, Lawn M. The origin of factor scores: Spearman, Thomson and Bartlett. *Br J Math Stat Psychol* 2009;**62**(3):569–82.
34. van Buuren S, Groothuis-Oudshoorn K. MICE: multivariate imputation by chained equations in R. *J Stat Software* 2011;**45**(3):1–67.
35. Clark TS, Linzer DA. Should I use fixed or random effects? *Polit Sci Res Methods* 2015;**3**(2):399–408.
36. Zhang X, Zhu W, Kang S, Qiu L, Lu Z, Sun Y. Association between physical activity and mood states of children and adolescents in social isolation during the COVID-19 epidemic. *J Environ Res* 2020;**17**(20):7666.
37. Reece LJ, Owen K, Foley B, Rose C, Bellew B, Bauman A. Understanding the impact of COVID-19 on children's physical activity levels in NSW, Australia. *J Aust* 2020 Nov. https://doi.org/10.1002/hpja.436.
38. Tulchin-Francis K, Stevens WJ, Gu X, Zhang T, Roberts H, Keller J, et al. The impact of the coronavirus disease 2019 pandemic on physical activity in US children. *J Sport Health Sci* 2021 May;**10**(3):323–32.
39. Medrano M, Cadenas-Sanchez C, Oses M, Arenaza L, Amasene M, Labayen I. Changes in lifestyle behaviours during the COVID-19 confinement in Spanish children: a longitudinal analysis from the MUGI project. *Pediatr Obes* 2021;**16**(4):e12731.
40. Xiang M, Zhang Z, Kuwahara K. Impact of COVID-19 pandemic on children and adolescents' lifestyle behavior larger than expected. *Prog Cardiovasc Dis* 2020 Aug;**63**(4):531–2.
41. Ng K, Cooper J, McHale F, Clifford J, Woods C. Barriers and facilitators to changes in adolescent physical activity during COVID-19. *BMJ Open Sport Exerc Med* 2020;**6**(1):e000919.
42. Maunula L. *Doing research in schools during the COVID-19 pandemic. What did we hear from students and parents in semi-structured interviews?* Canadian Public Health Association Forum; 2021 Apr 22. Online. Available from: https://www.cpha.ca/publichealth2021-program.
43. Adams EL, Caccavale LJ, Smith D, Bean MK. Food insecurity, the home food environment, and parent feeding practices in the era of COVID-19. *Obesity* 2020;**28**(11):2056–63.
44. Ruiz-Roso MB, de Carvalho Padilha P, Matilla-Escalante DC, Brun P, Ulloa N, Acevedo-Correa D, et al. Changes of physical activity and ultra-processed food consumption in adolescents from different countries during Covid-19 pandemic: an observational study. *Nutrients* 2020 Jul;**12**(8):2289.
45. Janssen M, Chang BPI, Hristov H, Pravst I, Profeta A, Millard J. Changes in food consumption during the COVID-19 pandemic: analysis of consumer survey data from the first lockdown period in Denmark, Germany, and Slovenia. *Front Nutr* 2021;**1**:635859.
46. Carroll N, Sadowski A, Laila A, Hruska V, Nixon M, Ma DWL, et al. The impact of COVID-19 on health behavior, stress, financial and food security among middle to high income Canadian families with young children. *Nutrients* 2020 Aug;**12**(8):2352.
47. Zhou S-J, Wang L-L, Yang R, Yang X-J, Zhang L-G, Guo Z-C, et al. Sleep problems among Chinese adolescents and young adults during the coronavirus-2019 pandemic. *Sleep Med* 2020;**1**:39–47.
48. Gruber R, Saha S, Somerville G, Boursier J, Wise MS. The impact of COVID-19 related school shutdown on sleep in adolescents: a natural experiment. *Sleep Med* 2020;**1**:33–5.
49. Lim MTC, Ramamurthy MB, Aishworiya R, Rajgor DD, Tran AP, Hiriyur P, et al. School closure during the coronavirus disease 2019 (COVID-19) pandemic – impact on children's sleep. *Sleep Med* 2021 Feb 1;**78**:108–14.
50. Bélanger RE, Patte KA, Leatherdale ST, Gansaonré RJ, Haddad S. An impact analysis of the early months of the COVID-19 pandemic on mental health in a prospective cohort of Canadian adolescents. *J Adolesc Health* 2021. https://doi.org/10.1016/j.jadohealth.2021.07.039.
51. Penner F, Ortiz JH, Sharp C. Change in youth mental health during the COVID-19 pandemic in a majority Hispanic/Latinx US sample. *J Am Acad Child Adolesc Psychiatry* 2021 Apr;**60**(4):513–23.
52. Choi J, Park Y, Kim H-E, Song J, Lee D, Lee E, et al. Daily life changes and life satisfaction among Korean school-aged children in the COVID-19 pandemic. *Int J Environ Res Publ Health* 2021 Mar 23;**18**:3324.
53. Kirmayer L, Sehdev M, Whitley R, Dandeneau S, Isaac C. Community resilience: models, metaphors and measures. *J Aborig Health* 2009:62–117.
54. Peters PA. Broadening the narrative on rural health: from disadvantage to resilience. *Univ Toronto Med J* 2019;**96**(2):41–3.
55. Vander Ploeg KA, McGavock J, Maximova K, Veugelers PJ. School-based health promotion and physical activity during and after school hours. *Pediatrics* 2014 Feb 1;**133**(2):e371–8.
56. Vander Ploeg KA, Maximova K, McGavock J, Davis W, Veugelers P. Do school-based physical activity interventions increase or reduce inequalities in health? *Soc Sci Med* 2014 Jul;**112**:80–7.
57. Bastian KA, Maximova K, McGavock J, Veugelers P. Does school-based health promotion affect physical activity on weekends? And, does it reach those students most in need of health promotion? Ozakinci G, editor. *PLOS ONE* 2015 Oct 21;**10**(10):e0137987.
58. Fung C, Kuhle S, Lu C, Purcell M, Schwartz M, Storey K, et al. From "best practice" to "next practice": the effectiveness of school-based health promotion in improving healthy eating and physical activity and preventing childhood obesity. *Int J Behav Nutr Phys Act* 2012;**9**(1):27.
59. Sithole F, Veugelers P. *Health reports: parent and child reports of children's activity*. 2015. Available from: https://www150.statcan.gc.ca/n1/pub/82-003-x/2008003/article/10662-eng.htm.
60. Munasinghe S, Sperandei S, Freebairn L, Conroy E, Jani H, Marjanovic S, et al. The impact of physical distancing policies during the COVID-19 pandemic on health and well-being among Australian adolescents. *J Adolesc Health* 2020 Nov;**67**(5):653–61.
61. Wright L, Steptoe A, Fancourt D. Are we all in this together? Longitudinal assessment of cumulative adversities by socioeconomic position in the first 3 weeks of lockdown in the UK. *J Epidemiol* 2020;**74**(9):683–8.

Public Health 202 (2022) 49–51

Contents lists available at ScienceDirect

Public Health

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

Short Communication

Practitioners' perspectives on health in Strategic Environmental Assessment of spatial planning policies in Scotland

C. Logue*, C. Werner, M. Douglas

Usher Institute – University of Edinburgh, Old Medical School, Teviot Place, Edinburgh, EH8 9AG, UK

ARTICLE INFO

Article history:
Received 21 December 2020
Received in revised form
4 October 2021
Accepted 18 October 2021
Available online 8 December 2021

Keywords:
Environmental assessment
Health impact assessment
Spatial planning

ABSTRACT

Objectives: Local authorities in Scotland are required to produce a Local Development Plan (LDP), which allocates sites for development and sets policies to guide decisions on planning applications. As part of this, local authorities must undertake a Strategic Environmental Assessment (SEA). This is a structured assessment of likely environmental impacts, which includes human health. This study explores how SEA practitioners and SEA consultation authorities consider health.
Study design: Qualitative study design using eight in-depth semi-structured interviews.
Methods: Individual interviews were carried out with SEA practitioners from six local authority areas in Scotland and two SEA consultation authorities. Interviews were recorded, transcribed and analysed thematically.
Results: Respondents articulated a broad perspective on health, but this was not reflected in SEA practice. Barriers to considering health more fully in SEA included low confidence in assessing health, limited partnership working with public health professionals and the lack of a consultation authority able to cover all aspects of health. Respondents valued partnership work between public health and planning professionals.
Conclusion: This study suggests recent work in Scotland to increase understanding of the role of spatial planning to influence health has been successful. However, further work is required to expand this to include links between spatial planning and health inequalities. SEA in Scotland does not currently support holistic consideration of health and health inequalities. Strong partnership working between public health and other sectors can increase understanding of links with health and create healthy places.

Spatial planning policy influences health because characteristics of the places where people live, work, grow and age affect their health and wellbeing. COVID-19 has highlighted further the importance of local environments to communities' wellbeing, including mental health impacts. In Scotland, public health and planning professionals increasingly recognise these links. The Public Health Priorities include 'A Scotland where we live in vibrant, healthy and safe places and communities'.[1] The Scottish National Planning Framework 4 (NPF4) Position Statement states 'NPF4 will be redesigned to support the population's health and wellbeing and address longstanding health inequalities'.[2] The relationship between place and health acts through multiple pathways. The Place Standard identifies fourteen dimensions, which are features of a place that affect health and may be influenced by spatial plans.[3] The Place Standard supports the engagement of communities to increase understanding of how a place affects their health.

All local authorities in Scotland must produce a Local Development Plan (LDP). This allocates sites for development and sets policies to guide planning decisions. A mandatory part of LDP development is Strategic Environmental Assessment (SEA), a structured assessment of likely environmental impacts. The Environmental Assessment (Scotland) Act 2005[4] establishes the issues to be considered in SEA, including human health.

The World Health Organisation advocates encompassing health impact assessments (HIA) within SEAs to ensure the full range of health determinants is considered.[5] Conversely, SEA guidance focuses narrowly on physical determinants. The Scottish

* Corresponding author. Screening and Early Detection Service, Canaan Park, Astley Ainslie Hospital, 143 Grange Loan, Edinburgh, EH9 2HL, UK. Tel.: +44 07787002454.
E-mail address: Ciara.logue@nhslothian.scot.nhs.uk (C. Logue).

https://doi.org/10.1016/j.puhe.2021.10.011

Environmental Protection Agency (SEPA) scrutinises human health in SEAs in Scotland. Its guidance directs assessors to consider 'environmentally related health issues such as exposure to traffic noise or air pollutants'.[6] One of us (MD) reviewed Scottish SEAs, including 15 SEAs of spatial plans, finding variation in the health impacts considered and little assessment of differential impacts.[7] Reviews in other settings similarly report that 'health' in SEAs focus narrowly on air, water and soil quality, as well as noise, with only a few considering differential impacts and health inequalities.[8,9] Most previous studies have been documentary reviews.

Little is currently known about the perspectives of SEA practitioners on including health in SEAs. We took a qualitative approach to explore the views of practitioners involved in SEA of LDPs in Scottish local authorities. Using qualitative semi-structured interviews, we explored the views of eight practitioners involved in the SEA of LDPs in Scottish local authorities. They were Planning Officers in six local authorities who had led SEAs of their authorities' LDPs and SEA managers in two Consultation Authorities who scrutinised SEA reports. We identified four main themes: (1) Broad concept of health and place (2) Narrow scope of health in SEA (3) Barriers to considering health in SEA (4) Strategies to improve health in SEA. Table 1 presents example quotes by theme.

Broad concept of health and place

All respondents described an understanding of health encompassing physical and mental wellbeing. They emphasised the breadth of this understanding, often using the words 'broad' or 'wide' and described multiple influences on health. Local authority respondents reported that their LDPs reflected this broad understanding and were designed to improve health. One interviewee described achieving this through working with health partners to embed health in planning.

Respondents noted recent increases in recognition of broad links between health and place. Five of the six local authority respondents described using the Place Standard to assess the quality of local places and links to health, which may have contributed to this. Several respondents identified poverty as a health determinant, but there was less recognition of how planning could influence health inequalities.

Narrow scope of health in SEA

Contrasting with their own broad understanding of health, most participants reported a narrower scope of health within SEA, including environmental hazards and greenspace but not mental or social wellbeing. None of the respondents reported consideration of differential impacts, although SEAs may include inequalities data in the baseline report. Health was viewed as subjective and less tangible than other issues assessed in SEA.

Several respondents reported frustration with SEA, saying it was cumbersome and did not add value to their planning process. They described LDPs as contributing to health improvement, reflecting good planning practice and Scottish Government policies. However, SEA did not enhance these opportunities.

There was one discrepant account from LA2, the only local authority respondent who was a dedicated SEA officer without a spatial planning background. LA2 argued that the whole purpose of SEA was to protect people's health through environmental improvements. Interestingly though, LA2 also recognised that SEA practice often excluded consideration of 'softer' health issues and suggested that involvement of health professionals could help address this.

Barriers to considering health in SEA

Challenges in assessing health in SEAs included confidence and expertise of SEA practitioners. Some mentioned loss of dedicated SEA officers due to savings. Most practitioners reported lacking confidence in assessing health issues. There were difficulties in engaging with NHS colleagues in SEA, although some had engaged with health colleagues in developing the LDP itself. Several respondents identified a lack of relevant health data or expertise to understand and use health data meaningfully. Finally, they identified a lack of independent scrutiny of health within SEAs. Each consultation authority scrutinises SEAs in relation to their area of

Table 1
Interviewee quotes by theme.

Theme	Quote	Interviewee[a]
1) Broad concept of health and place	*'I suppose within planning there's been a real kind of raising of awareness in the last 4 or 5 years, of the role that planning has in terms of creating places but it's also how that physical environment links quite dramatically with people's health … access to health services … active travel – walking, cycling opportunities, how a new development is going to affect the number of cars on the road, then that feeds into air quality. So, I think in my time here I have seen a real kind of, almost a change in mindset, that planning's not just the physical environment it actually affects peoples physical and mental health as well.'*	LA4
2) Narrow scope of health in SEA	*'When I put my SEA hat on that just narrows everything right down.. our [LDP] objective in terms of health is provide a suitable range of housing and employment opportunities, improve the health and living environment of people and communities, so they are really quite broad. Then the [SEA] questions are … how will it affect people in terms of noise or smells etc and do they have access to cycling and walking routes, open spaces and green spaces.'*	LA4
(Discrepant account)	*'The fundamental environmental thing for SEA is people and their health, and that's why in most things we do in terms of SEA, that is always top consideration, is people and health. As no matter what environmental consideration you look at in terms of SEA, it all comes back to the impact on or effect it would have on people, and their health and their wellbeing … like the green thread.'*	LA2
3) Barriers to considering health in SEA	*'They struggle to find out what is relevant health information in the context of the plan they are preparing so there is a big challenge there. They couldn't find information, even when they did have information they didn't really know what to do with it.'*	CA1
4) Strategies to improve health in SEA	*'I think [health] guidance would only be useful if you had, if the health agenda was picked up more clearly in one of the consultation authority's roles … unless you have someone in a position to pick up that aspect of SEA it's probably going to sit on the shelf somewhere.'*	CA2
	'[in workshops] … everyone talks to each other about all the different things and how it impacts. So now when we do environmental work here, people aren't sitting in silos doing their own thing, it tends to be more joined up.'	LA2

[a] LA indicates a Local Authority interviewee, and CA indicates a Consultation Authority interviewee.

expertise and may identify some impacts relevant to health, but they cannot assess health overall.

Strategies to improve coverage of health in SEA

Several respondents proposed specific guidance or a template to identify health issues in SEA. This could build on tools used in health impact assessments adapted for SEA. However, others argued for a more integrated, streamlined approach. Some reflected on whether the Place Standard could be used in SEA, though there were some concerns that it was too 'subjective'. Several reported positive experiences using it for local neighbourhoods but felt it would be less useful at a higher scale.

Respondents agreed that health would be more fully covered in SEA if there were a consultation authority able to scrutinise the full range of health issues, although most were pessimistic about identifying a suitable authority to take on this role.

Several respondents identified potential benefits of greater partnership work with health colleagues. LA2, who expressed the most enthusiasm for SEA and its potential to improve health, was the only practitioner who reported good involvement of NHS colleagues in SEA. This involved workshops at the scoping stage of SEA.

Embedding health in planning – the way forward

The accounts from SEA practitioners reported within this paper demonstrate their understanding of health, as well as their perceptions of SEA. Practitioners in our study were familiar with the Place Standard, which may have contributed to their understanding of links between health and spatial planning. This did not transfer, however, to a greater understanding of links with health inequalities. Further work should raise awareness of the equity impacts of planning.

The WHO has long advocated the inclusion of health within SEA,[5] but SEA remains narrowly focused on environmental risks, despite repeated arguments for change. The structure of SEA also provides a poor framework to assess differential impacts and health inequalities. This highlights a glaring question. Should public health professionals abandon efforts to influence SEA and instead engage with planning, and other sectors, in other ways? This could involve separate health impact assessments, tools like the Place Standard, shared data, evidence and further training. Public Health has scarce resources, with increased capacity constraints following COVID-19, so it must focus efforts in ways that are most likely to be influential. However, increased understanding of health among spatial planners in Scotland may not be replicated in other sectors, in settings where SEA guidance is less restrictive, or among the private sector consultants who complete many SEAs. As a mandatory assessment that is intended to include health, SEA should be viewed as a key opportunity to improve the health impact of plans and policies in many sectors. Doing this requires us to challenge guidance restricting the scope of health.[6] Recent international guidance may help to support this.[10]

Our findings highlight the important role of SEA consultation authorities in directing the scope and scrutinising SEAs. Respondents reported that Scotland lacks a consultation authority able to take a broad overview of health determinants in SEAs. Public Health Scotland is a newly established organisation covering all the domains of public health and could take on this role, but would need dedicated resources and legislative support to do so.

Finally, our findings highlight the benefits of a strong partnership. The accounts suggest the Place Standard, developed as a partnership between public health and planning, has broadened the understanding of health among planners. Respondents also reported benefits from involving public health colleagues, either in LDP development or less commonly in the SEA. The health in all policies approach involves public health professionals building working relationships with colleagues to influence plans and policies. This can involve working together on SEA or through other processes. Either way, partnerships and public health engagement in planning are a powerful way to create a common understanding of opportunities for health and contribute to healthier policies in planning and other sectors.

Author statements

Ethical approval

Level 2 ethical approval was granted for this research by Usher Research Ethics Group at University of Edinburgh.

Funding

None declared.

Competing interests

None declared.

References

1. Scottish Government. *Public Health Priorities for Scotland*. 2018. https://www.gov.scot/publications/scotlands-public-health-priorities/. [Accessed 1 July 2021].
2. Scottish Government. *Fourth National Planning Framework: Position Statement*. 2020. https://www.gov.scot/publications/scotlands-fourth-national-planning-framework-position-statement/pages/5/. [Accessed 1 July 2021].
3. Scottish Government. *Place Standard Tool Strategic Plan 2020-2023*. 2020. https://www.placestandard.scot/docs/Place_Standard_Strategic_Plan.pdf. [Accessed 1 July 2021].
4. Scottish Government. *Environmental Assessment (Scotland) Act*. 2005. https://www.legislation.gov.uk/asp/2005/15/contents. [Accessed 1 July 2021].
5. Fehr R, Viliani F, Nowacki J, Martuzzi M, editors. *Health in Impact Assessments: Opportunities not to be Missed*. Copenhagen: WHO Regional Office for Europe; 2014. http://www.euro.who.int/en/health-topics/environment-and-health/health-impact-assessment/publications/2014/health-in-impact-assessments-opportunities-not-to-be-missed. [Accessed 1 July 2021].
6. Scottish Environment Protection Agency. *Strategic Environmental Assessment SEPA Guidance Note 5: Guidance on Consideration of Human Health in Strategic Environmental Assessment*. 2019. https://www.sepa.org.uk/media/219433/lups-sea-gu5-consideration-of-human-health-in-sea.pdf. [Accessed 1 July 2021].
7. Douglas M, Carver H, Katikireddi S. How well do strategic environmental assessments in Scotland consider human health? *Public Health* 2011;**125**(9):585–91. https://www.sciencedirect.com/science/article/pii/S0033350611001892.
8. Fischer T, Matuzzi M, Nowacki J. The consideration of health in strategic environmental assessment (SEA). *Environ Impact Assess Rev* 2010;**30**(3):200–10. https://www.sciencedirect.com/science/article/pii/S0195925509001292.
9. Carmichael L, Barton H, Gray S, Lease H, Pilkington P. Integration of health into urban spatial planning through impact assessment: identifying governance and policy barriers and facilitators. *Environ Impact Assess Rev* 2012;**32**(1):187–94.
10. Economic Commission for Europe. *Draft Guidance on Assessing Health Impacts in Strategic Environmental Assessment*. 2020. ECE/MP.EIA/WG.2/2020/7, https://unece.org/fileadmin/DAM/env/eia/documents/WG.9_2020/Final_documents/2004508E.pdf. [Accessed 1 July 2021].

Public Health 202 (2022) 106–112

Contents lists available at ScienceDirect

Public Health

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

Original Research

Reactions to geographic data visualization of infectious disease outbreaks: an experiment on the effectiveness of data presentation format and past occurrence information

J. Zhang [a], Y. Wang [b], W. Wanta [c], Q. Zheng [d], X. Wang [e, *]

[a] Polk School of Communications, Long Island University, Brooklyn, NY, USA
[b] Department of Communication, University of Louisville, Louisville, KY, USA
[c] College of Journalism and Communications, University of Florida, Gainesville, FL, USA
[d] School of Public Health & Information Sciences, University of Louisville, Louisville, KY, USA
[e] School of New Media, Peking University, Beijing, China

ARTICLE INFO

Article history:
Received 10 June 2021
Received in revised form
26 September 2021
Accepted 14 November 2021
Available online 20 December 2021

Keywords:
Data visualization
Thematic map
Comprehension
Memory
Risk communication
Zika

ABSTRACT

Objectives: This study intended to compare the effectiveness of thematic maps with that of tabular data in comprehension and memory of risk magnitudes, with Zika virus (ZIKV) disease outbreaks in the United States as the subject matter. The study also aimed to examine the effects of data presentation format and past occurrence information on risk perception and risk avoidance intention.
Study design: This study used an experiment.
Methods: Each participant was randomly assigned to view ZIKV disease 2017 incidence data presented in one of the three formats: a choropleth map, a graduated-circle map, and a table, after which they answered questions about comprehension and memory of risk magnitudes. Each participant was then randomly assigned to view or not to view incidence data of the previous occurrence of ZIKV outbreaks in 2016, after which they answered questions about risk perception and risk avoidance intention.
Results: The results revealed the effectiveness of thematic maps over tabular data in comprehension, risk perception, and risk avoidance intention. Compared to tabular data, the choropleth map led to a better comprehension of relative risk magnitudes, the graduated-circle map led to higher risk perception, and both thematic maps led to greater risk avoidance intention. In contrast, tabular data led to better recognition of absolute risk magnitudes than both thematic maps. In addition, past occurrence information enhanced risk perception and risk avoidance intention.
Conclusions: The findings reveal the importance of data presentation format in comprehension and memory of risk magnitudes. This can be attributed to the cognitive match between the information emphasized in the presentation and that required by the tasks. The findings also suggest that data presentation format and past occurrence information are important judgmental heuristics that help to form risk perception and risk avoidance intention.

Introduction

Using geographic data visualization to convey disease outbreak data, which could go back as early as Dr. John Snow's mapping of the 1854 London cholera outbreak,[1] has become more common in recent years due to the development of computer technology. With ZIKV disease outbreaks in the United States as the subject matter, this study intended to compare the effectiveness of two types of thematic maps—choropleth maps and graduated-circle maps—with that of tabular data in comprehension and memory of risk magnitudes. This study also aimed to examine the impact of data presentation format and past occurrence information on risk perception and risk avoidance intention. The findings provided implications of how best to present data to effectively communicate disease outbreak information to the public.

* Corresponding author.
E-mail address: xiuli.wang@pku.edu.cn (X. Wang).

https://doi.org/10.1016/j.puhe.2021.11.011

Literature review

The effects of data presentation format on comprehension and memory of risk magnitudes

Tables have been conventionally used to display the incidence of a disease or the number of new cases of a disease.[2] In recent years, thematic maps have been increasingly used to visualize disease outbreak data. Two of the commonly used maps are choropleth maps that symbolize data with color shading and graduated-circle maps that symbolize data with proportional circles.[1,3,4]

The effectiveness of the data presentation format has been examined by measures such as comprehension and memory of health risk information.[5,6] Comprehension can be assessed by the subjective easiness of understanding the information presented.[6] Memory is assessed by recognition and recall.[7] Recognition is about recognizing a previously encountered stimulus. The recall involves retrieving information despite the absence of a stimulus,[7,8] which includes cued recall with cues and free recall without cues. Generally, recognition is the easiest, and free recall is the hardest.[8,9]

In health risk research, there has been a distinction between absolute and relative risk magnitudes.[5,10,11] Absolute risk magnitudes are measured by extracting discrete data values,[12,13] whereas relative risk magnitudes are gauged by making comparisons and perceiving relationships in the data.[1,3,14,15] According to the cognitive fit theory, a match compared to a mismatch between information emphasized in the presentation and that required by a task enhances task performance.[12,16–18] Thematic maps, which emphasize geographic patterns in the data, tend to facilitate tasks assessing relative risk magnitudes. Therefore, we predicted that thematic maps would lead to greater comprehension and better free recall of relative risk magnitudes than tabular data.

Hypotheses 1. (H1): Participants presented with thematic maps will find it easier to (a) understand the condition of ZIKV disease outbreaks in the United States, and (b) understand the difference in the ZIKV disease outbreaks by state and territory than participants presented with tabular data.

Hypothesis 2. (H2): Participants presented with thematic maps will have better free recall of high-risk states and territories with the greatest number of ZIKV disease cases than participants presented with tabular data.

Tables, which emphasize discrete data values, tend to enhance tasks assessing absolute risk magnitudes. Therefore, we predicted that tabular data would lead to better recognition of absolute risk magnitudes than thematic maps.

Hypothesis 3. (H3): Participants presented with tabular data will have better recognition of (a) the number of ZIKV disease cases for the state or territory they primarily lived in and (b) the number of ZIKV disease cases for the most visited states, than participants presented with thematic maps.

The effects of data presentation format and past occurrence information on risk perception and risk avoidance intention

Judgmental heuristics such as information presentation and availability help people to form risk perception and make decisions under uncertainty.[12,19] Prior research has shown that graphical compared to numerical presentation increases risk avoidance.[5,20] One rationale is that graphical presentation heightens the cognitive impression of riskiness.[21] Geographic data visualization adds a geospatial context to the data and allows for the quick and joint investigation of statistical and geographic patterns in data.[3,15] Thus, we predicted that thematic maps would result in greater risk perception and risk avoidance intention than tabular data.

Hypothesis 4. (H4): Participants presented with thematic maps will perceive ZIKV as posing more of a threat to themselves than participants presented with tabular data.

Hypothesis 5. (H5): Participants presented with thematic maps will be more likely to (a) take steps to reduce the risk posed by the ZIKV and (b) avoid travel to high-risk areas for ZIKV transmission than participants presented with tabular data.

Information about the relevant occurrence of health risks has been shown to increase risk perception.[21,22] Information about the previous occurrence of ZIKV disease outbreaks may enhance its salience, thus functioning as a mental shortcut for risk judgment.[20] We collected data during the peak summer season of the spread of ZIKV in 2017 with the uncertainty of what the scale would finally be. As ZIKV outbreaks were notifiable in the United States in the previous year 2016,[23] additional presentation of incidence data from 2016 may increase risk perception and risk avoidance intention in the year 2017. Thus, we predicted that past occurrence information would facilitate risk perception and risk avoidance intention.

Hypothesis 6. (H6): Past occurrence information will enhance risk perception.

Hypothesis 7. (H7): Past occurrence information will enhance the intention to (a) take steps to reduce the risk posed by the ZIKV and (b) avoid travel to high-risk areas for ZIKV transmission.

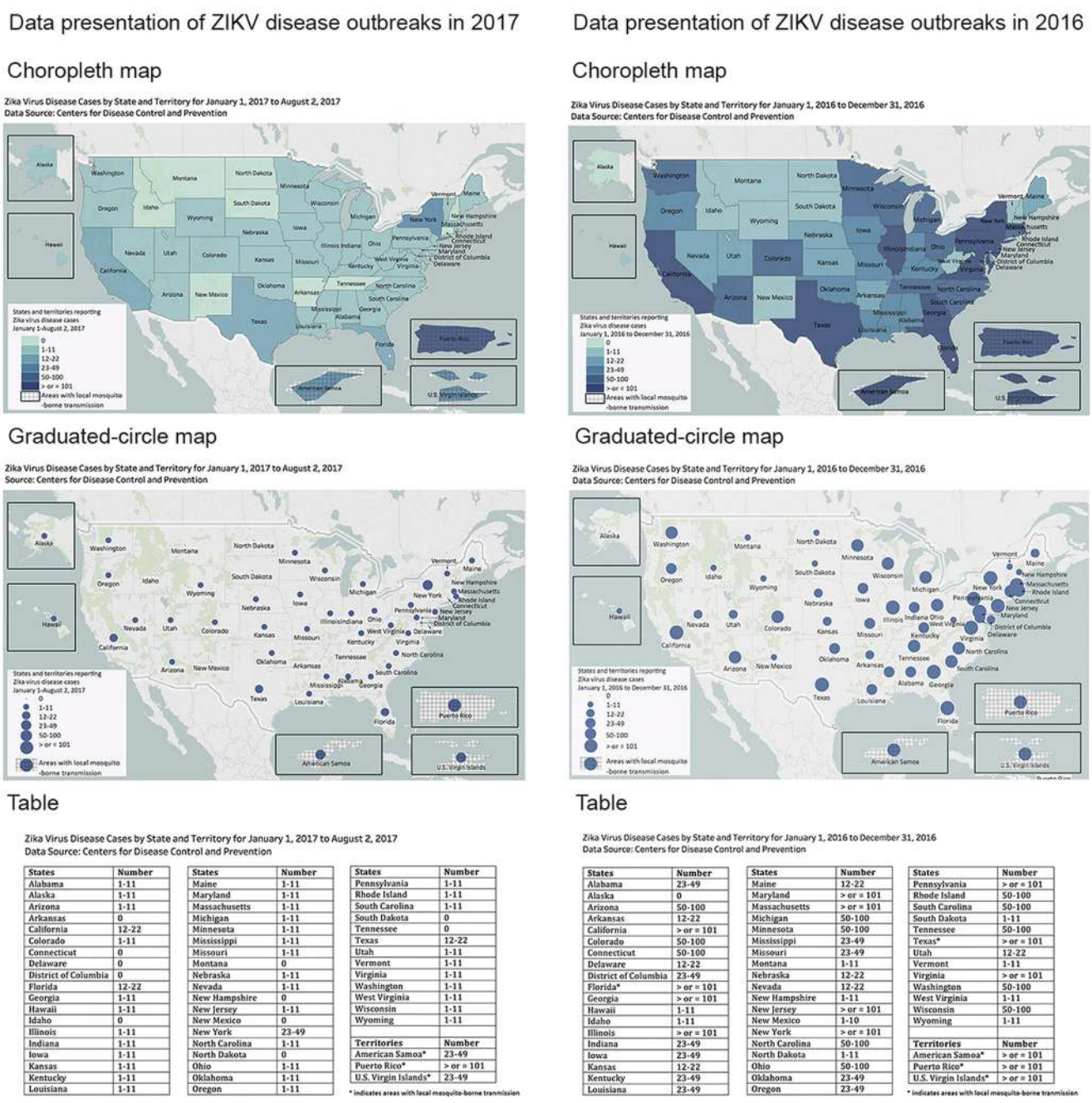
Methods

Participants

Participants were recruited using the Amazon Mechanical Turk (MTurk) online service in August 2017. Participants were at least 18 years old to participate in the study. Other criteria were: (1) The participant was in the United States; (2) The participant's Human Intelligence Task (HIT) approval rate for all requesters' HITs was greater than 95 out of 100; and (3) The number of HITs approved was greater than 100. A HIT is a task created by a requester for a participant to complete on MTurk.[24]

Stimuli

We created the stimuli based on the Centers for Disease Control and Prevention's (CDC) data of ZIKV disease cases in the United States.[23] Choropleth maps and graduated-circle maps were created using Tableau, and the tables were created using Microsoft Word. The CDC classified the numbers of cases into six ordinal classes: (1) 0, (2) 1–11, (3) 12–22, (4) 23–49, (5) 50–100, and (6) > or = 101. Prior research suggests that an effective classification for thematic maps is to define four to six classes.[25] Accordingly, we adopted the CDC's classification for the stimuli. Choropleth maps use a single-hue sequential shading scheme to represent ordinal classes by the spatial unit of states or territories with darker shading representing higher incidence, whereas graduated-circle maps use a range of circles with larger circles representing higher incidence.[3,26] Also, the color blue was chosen for the thematic maps to avoid eliciting emotional responses.[25] Please see Fig. 1 for the stimuli.



States	Number
Alabama	1-11
Alaska	1-11
Arizona	1-11
Arkansas	0
California	12-22
Colorado	1-11
Connecticut	0
Delaware	0
District of Columbia	0
Florida	12-22
Georgia	1-11
Hawaii	1-11
Idaho	0
Illinois	1-11
Indiana	1-11
Iowa	1-11
Kansas	1-11
Kentucky	1-11
Louisiana	1-11

States	Number
Maine	1-11
Maryland	1-11
Massachusetts	1-11
Michigan	1-11
Minnesota	1-11
Mississippi	1-11
Missouri	1-11
Montana	0
Nebraska	1-11
Nevada	1-11
New Hampshire	0
New Jersey	1-11
New Mexico	0
New York	23-49
North Carolina	1-11
North Dakota	0
Ohio	1-11
Oklahoma	1-11
Oregon	1-11

States	Number
Pennsylvania	1-11
Rhode Island	1-11
South Carolina	1-11
South Dakota	0
Tennessee	0
Texas	12-22
Utah	1-11
Vermont	1-11
Virginia	1-11
Washington	1-11
West Virginia	1-11
Wisconsin	1-11
Wyoming	1-11
Territories	**Number**
American Samoa*	23-49
Puerto Rico*	> or = 101
U.S. Virgin Islands*	23-49

* indicates areas with local mosquito-borne tranmission

States	Number
Alabama	23-49
Alaska	0
Arizona	50-100
Arkansas	12-22
California	> or = 101
Colorado	50-100
Connecticut	50-100
Delaware	12-22
District of Columbia	23-49
Florida*	> or = 101
Georgia	> or = 101
Hawaii	1-11
Idaho	1-11
Illinois	> or = 101
Indiana	23-49
Iowa	23-49
Kansas	12-22
Kentucky	23-49
Louisiana	23-49

States	Number
Maine	12-22
Maryland	> or = 101
Massachusetts	> or = 101
Michigan	50-100
Minnesota	50-100
Mississippi	23-49
Missouri	23-49
Montana	1-11
Nebraska	12-22
Nevada	12-22
New Hampshire	1-11
New Jersey	> or = 101
New Mexico	1-10
New York	> or = 101
North Carolina	50-100
North Dakota	1-11
Ohio	50-100
Oklahoma	23-49
Oregon	23-49

States	Number
Pennsylvania	> or = 101
Rhode Island	50-100
South Carolina	50-100
South Dakota	1-11
Tennessee	50-100
Texas*	> or = 101
Utah	12-22
Vermont	1-11
Virginia	> or = 101
Washington	50-100
West Virginia	1-11
Wisconsin	50-100
Wyoming	1-11
Territories	**Number**
American Samoa*	> or = 101
Puerto Rico*	> or = 101
U.S. Virgin Islands*	> or = 101

* indicates areas with local mosquito-borne tranmission

Fig. 1. Data presentation of ZIKV disease outbreaks.

Procedure

Using the Qualtrics survey platform, each participant was randomly assigned to view the ZIKV disease incidence data in the United States from January 1, 2017, to August 2, 2017, presented in one of the three formats: a choropleth map, a graduated-circle map, or a table, after which, they answered questions assessing comprehension and memory. Each participant was then randomly assigned to view or not to view the ZIKV disease incidence data in the United States from January 1, 2016, to December 31, 2016. Those who viewed the 2016 data viewed it in the same format as they viewed the 2017 data. Next, all participants answered questions about risk perception and risk avoidance intention. Finally, participants answered basic demographic questions.

Measures

Comprehension of risk magnitudes

Comprehension of risk magnitudes was assessed in two aspects: (1) comprehension of the condition of outbreaks in the United States and (2) comprehension of the difference in outbreaks by state and territory. Comprehension of the condition of outbreaks in the United States was measured by a statement adapted from the previous research:[5,12] 'The information makes it easy to understand

the condition of Zika virus disease outbreaks in the US.' Comprehension of difference in outbreaks by state and territory was measured by two statements adapted from prior research:[10] 'The information makes it easy to understand which states and territories had the greatest number of Zika virus disease cases,' and 'The information makes it easy to compare the number of Zika virus disease cases by state and territory.' Ratings on the two statements were averaged into an index (Cronbach's alpha = .81). Answers were rated using a five-point Likert scale from 'strongly disagree' to 'strongly agree'.

Free recall of relative risk magnitudes

Free recall requires retrieval of information in any order.[8] Consistent with prior research that involved data comparison for spatial tasks of recall,[12] participants were asked to recall and write down in any order four states and three territories that 'had reported the greatest number of Zika virus disease cases'. We coded each correct answer as 1 point. Participants' scores were calculated based on the number of points they received. As they were asked to write down a total of seven high-risk states and territories, the range of the scores was from 0 to 7.

Recognition of absolute risk magnitudes

Previous research often assessed recognition using a forced-choice recognition test, which requires recognizing information previously presented by choosing among a finite number of alternatives.[8] Following the convention, participants were asked to select the response option that best described the absolute risk magnitude in the state or territory they primarily lived in, as well as in the five most-visited states in the United States, which were California, Florida, Nevada, Texas, and New York.[27] The response options for absolute risk magnitudes, which were consistent with the data value classes in the stimuli, were: (1) 0, (2) 1–11, (3) 12–22, (4) 23–49, (5) 50–100, and (6) > or = 101. We scored correctness based on how correct the answers were. A correct answer received 6 points. One point was deducted for one class away from the correct answer. Thus, the least correct answer, which could be five classes away from the correct one, received 1 point. For recognition of risk magnitudes of the five most visited states, a mean correctness score was calculated.

Risk perception

Risk perception was measured by two statements adapted from earlier research,[28] which assessed the extent to which participants thought 'the Zika virus posed a threat to themselves,' and 'the Zika virus posed a threat to the general public in the US.' Answers were rated using a five-point Likert scale from 'not at all' to 'a great deal'. Ratings on the two statements were averaged into an index (Cronbach's alpha = .78).

Risk avoidance intention

Participants rated on a five-point Likert scale from 'not at all' to 'a great deal' the extent to which they should (1) 'take steps to reduce the risk posed by Zika virus,' and (2) 'avoid travel to high-risk areas for Zika virus transmission,' which were preventive measures suggested by CDC.[23]

Control variables

Prior knowledge, interest, and relevance were treated as control variables. Before viewing the stimuli, participants rated on a five-point Likert scale from 'not at all' to 'a great deal' the extent to which: (1) they had heard about Zika virus, (2) they were interested in information about Zika virus, and (3) Zika virus was relevant to them. The statements were adapted from prior research.[28]

Demographics

Participants were asked general demographic questions, including age, sex, race and ethnicity, and education.

Data analysis

The univariate analysis of variance (ANOVA) was used to assess the effects of data presentation format on comprehension and memory. It was also used to examine the effects of data presentation format and prior occurrence information on risk perception and risk avoidance intention. Fisher's Least Significant Difference (LSD) was used for posthoc pairwise comparisons. In all the analyses, prior knowledge, interest, and relevance were treated as control variables.

Results

Participants

A total of 300 participants with an average age of 39 completed the study. A majority of the participants (180 or 60%) were female; 119 or 39.67% were male; and 1 or 0.33% was other. A majority of the participants (214 or 71.33%) were White; 29 or 9.67% were Black; 16 or 5.33% were Hispanic; 26 or 8.67% were Asian; 1 or 0.33% was Native American; 2 or 0.67% were other, and 12 or 4.00% reported two races. A majority of the participants (126 or 42%) had bachelor's degrees; 60 or 20% had some college without a degree; 41 or 13.67% had associate degrees; 41 or 13.67% had master's degrees; 23 or 7.67% had high school degrees; 4 or 1.33% had doctorate degrees; 3 or 1.00% had professional degrees; and 2 or 0.07% had less than a high school diploma.

Before viewing the stimuli, the extent to which participants had heard about the ZIKV averaged 3.52 (SD = .86); that they were interested in information about ZIKV averaged 3.27 (SD = 1.03); and that ZIKV was relevant to them averaged 2.78 (SD = 1.12).

Effects of data presentation format on comprehension and memory

Table 1 illustrates the descriptive information and statistics about the effects of data presentation format on comprehension and memory.

Data presentation affected the easiness to understand the condition of ZIKV disease outbreaks in the United States, F (2, 294) = 3.359, P = .036, $\eta 2$ = .022. Fisher's LSD posthoc pairwise comparisons indicated the advantage of the choropleth map over the table (P = .011), but there was no difference between the graduated-circle map and the table. Data presentation also differentiated the easiness to understand the difference in risk magnitudes by state and territory, F (2, 294) = 7.701, P = .001, $\eta 2$ = .050. The choropleth map outperformed the table (P = .032), but there was no difference between the graduated-circle map and the table. H1 was partially supported, as the choropleth map, but not the graduated-circle map, was easier for risk comprehension than the table.

H2 predicted the superiority of thematic maps over tabular data in free recall of high-risk states and territories. H2 was not supported.

Data presentation affected recognition of ZIKV disease incidence counts for the state or territory participants primarily lived in, F (2, 294) = 8.530, P < .001, $\eta 2$ = .055, with the table outperforming both the choropleth map (P = .015) and the graduated-circle map (P < .001). Data presentation also affected recognition of ZIKV disease incidence counts of the five most-visited states, F (2, 294) = 31.970, P < .001, $\eta 2$ = .179; with the table outperforming both

Table 1
Effects of data presentation format on comprehension and memory.

Dependent Variables	Choropleth Map	Graduated-circle map	Tabular data	F	df	*P*	Partial eta squared
	Mean (SD)	Mean (SD)	Mean (SD)				
Comprehension							
Easy to understand the condition of ZIKV disease outbreaks in the US	4.25 (.71)	4.13 (.75)	4.01 (.90)	3.359	2,294	.036	.022
Easy to understand the difference of risk magnitudes by state and territory	4.29 (.67)	3.85 (.94)	4.05 (.91)	7.701	2,294	.001	.050
Memory							
Free recall of high-risk states and territories	5.21 (1.55)	5.31 (1.73)	5.54 (1.45)	.739	2,294	*n.s.*	
Recognition of ZIKV disease incidence counts for the state or territory participants primarily lived in	5.52 (.78)	5.31 (.82)	5.77 (.73)	8.530	2,294	< .001	.055
Recognition of ZIKV disease incidence counts in the five most visited states	5.10 (.63)	4.86 (.67)	5.55 (.51)	31.970	2,294	< .001	.179

the choropleth map ($P < .001$) and the graduated-circle map ($P < .001$). H3 was supported.

Effects of data presentation format and past occurrence information on risk perception and risk avoidance intention

Table 2 reveals the descriptive information and statistics about the effects of data presentation format and past occurrence information on risk perception.

Data presentation affected risk perception, $F\ (2, 291) = 5.133$, $P = .006$, $\eta 2 = .034$. The superiority of the graduated-circle map over the table was significant ($P = .002$), but not that of the choropleth map over the table. H4 was partially supported, as the graduated-circle map, but not the choropleth map, led to greater risk perception than the table.

Participants presented with incidence data of the past occurrence of ZIKV outbreaks regarded ZIKV as posing more of a threat, $F\ (1, 291) = 13.598$, $P < .001$, $\eta 2 = .045$. H6 was supported.

Table 3 shows the descriptive information and statistics about the effects of data presentation format and past occurrence information on risk avoidance intention.

Data presentation affected the intention to take steps to reduce the risk posed by ZIKV, $F\ (2, 291) = 7.156$, $P = .001$, $\eta 2 = .047$. Both the choropleth map and the graduated-circle map outperformed the table with $P = .002$ and $P = .001$ respectively. Also, data presentation differentiated the intention to avoid travel to high-risk areas for ZIKV transmission, $F\ (2, 291) = 5.601$, $P = .004$, $\eta 2 = .037$. Both the choropleth map and the graduated-circle map outperformed the table with $P = .003$ and $P = .005$ respectively. H5 was supported.

H7 predicted that prior occurrence information would increase risk avoidance intention. Participants presented with past occurrence information were more likely to take steps to reduce the risk posed by ZIKV, $F\ (1, 291) = 7.637$, $P = .006$, $\eta 2 = .026$; but not to avoid travel to high-risk areas for ZIKV transmission. H7 was partially supported, as prior occurrence information increased the intention to take preventive steps generally, but not to avoid travel to high-risk areas.

There was no interaction effect between data presentation format and past occurrence information on risk perception and risk avoidance intention.

Discussion

The first purpose of this study was to examine the impact of data presentation format on comprehension and memory of risk magnitudes. According to the cognitive fit theory, thematic maps emphasize geographic patterns in the data and thus tend to facilitate spatial oriented tasks involving comparison of risk magnitudes.[1,3,12,14,15] The findings revealed the advantage of the choropleth map, but not the graduated-circle map, over the table in both comprehension measures: the easiness to understand the condition of ZIKV disease outbreaks in the United States and the difference in the ZIKV disease outbreaks by state and territory. The findings suggest that the two types of thematic maps vary in their impact on comprehension. There are several plausible factors worth investigating. For example, the appearance of thematic maps is affected by arbitrary choices to categorize ordinal classes. Also, viewers' ability to associate symbolizations with ordinal classes may affect comprehension.[3]

In comparison, tabular data emphasize discrete data values and thus tend to enhance symbolic orientated tasks that involve extracting specific data values.[12,13] As predicted, the results revealed the superiority of tabular data over both thematic maps on both recognition tasks: correctness in memorizing the ZIKV incidence counts of the participant's primary residential state or territory and those of the five most-visited states.

As to free recall of relative risk magnitudes, there was no main effect of the data presentation format. It is worth noting that the average correctness score for free recall was 5.34 ($SD = 1.58$) out of

Table 2
Effects of data presentation format and past occurrence information on risk perception.

Main effects	Risk perception				
	Mean (SD)	F	df	*P*	Partial eta squared
Data presentation		5.133	2,291	.006	.034
Choropleth map	2.66 (1.05)				
Graduated-circle map	2.94 (.89)				
Tabular data	2.63 (.98)				
Past occurrence information		13.598	1,291	<.001	.045
Yes	2.87 (1.05)				
No	2.60 (.91)				

Table 3
Effects of data presentation format and past occurrence information on risk avoidance intention.

Main effects	Intention to take steps to reduce the risk					Intention to avoid travel to high-risk areas				
	Mean (SD)	F	df	*P*	Partial eta squared	Mean (SD)	F	df	*P*	Partial eta squared
Data presentation		7.156	2,291	.001	.047		5.601	2,291	.004	.037
Choropleth map	2.96 (1.25)					3.32 (1.40)				
Graduated-circle map	3.12(1.08)					3.39 (1.24)				
Tabular data	2.68 (1.16)					2.90 (1.35)				
Past occurrence information		7.637	1,291	.006	.026		2.910	1,291	n.s.	
Yes	3.04 (1.24)					3.31 (1.33)				
No	2.80 (1.11)					3.10 (1.37)				

7, which indicated good performance, although free recall could be the hardest retrieval task.[8,9] A couple of factors may have contributed to the good performance. First, ZIKV tended to spread faster in the states with larger populations and warmer weather, which may affect the free recall regardless of the data presentation format. Second, the three territories were grouped together and separated from the states in the data presentation, which may make them much easier to be memorized.

The second purpose of this study was to examine the impact of data presentation format and past occurrence information on risk perception and risk avoidance intention. Like other graphical presentations,[5,20] both thematic maps led to greater risk avoidance intention than did tabular data. One plausible rationale is that graphical presentation heightens the cognitive impression of riskiness.[21] In addition, past occurrence information enhanced risk perception. It also increased the intention to take steps to reduce the risk posed by ZIKV but not to avoid travel to high-risk areas for ZIKV transmission. It is plausible that additional information is needed to enhance intention to avoid travel to high-risk areas.

Implications for health risk communication

The choice of data presentation format depends on the goals of health risk communication. The choropleth map is most effective in enhancing comprehension of risk magnitudes, whereas the graduated-circle map is most effective in increasing risk perception. Both thematic maps are better than tabular data in the forming of risk avoidance intention. In comparison, tabular data is more effective than thematic maps in recognition of risk magnitudes of certain spatial units. Finally, adding past occurrence information will help to increase risk perception and risk avoidance intention.

Limitations and future studies

This study has several limitations. The sample was limited to Amazon MTurk online workers. The second limitation was that a single statement was used to measure the dependent variable of comprehension of the condition of outbreaks in the United States. Although single-item measures are arguably adequate for simple constructs, more pertinent statements could be used in future studies for plausibly higher validity.[29] Also, single-item scales were used to measure control variables. Future studies could use multi-item scales for probably high validity. Third, among the three data presentation formats, the graduated-circle map was the least effective in comprehension and recognition. It would be interesting for future research to examine if any characteristic of the graduated-circle map can provide explanations. Finally, inconsistent with the prediction, there was no advantage of thematic maps over the table in free recall of relative risk magnitudes. This could be due to the characteristics of high-risk states regardless of data presentation and the prominence of high-risk territories in the data presentated. It is worth further investigation.

Conclusions

The findings suggest the importance of data presentation format in comprehension and memory of risk magnitudes. The results indicate the superiority of the choropleth map over the table in comprehension of risk magnitudes and that of tabular data over both thematic maps in recognition of absolute risk magnitudes. This could be attributed to the cognitive match between the information emphasized in the presentation and that required by the tasks. The findings also suggest that data presentation format and past occurrence information are important judgmental heuristics that help to form risk perception and risk avoidance intention. There was an advantage of the graduated-circle map over the table in the forming of risk perception and that of both thematic maps over the table in the forming of risk avoidance intention. In addition, past occurrence information enhanced risk perception and risk avoidance intention.

Author statements

Ethics approval

The study received IRB approval from the New York Institute of Technology.

Funding

None.

Competing interests

None declared.

Author contributions

JZ and YW conceived the study, prepared data collection and analyses, and wrote the first draft of the manuscript. WW, QZ, and XW contributed to data analyses and provided important comments on the manuscript. All authors read and approved the final manuscript.

References

1. Smith CM, Le Comber SC, Fry H, Bull M, Leach S, Hayward AC. Spatial methods for infectious disease outbreak investigations: systematic literature review. *Euro Surveill* 2015;**20**(39):30026. https://doi.org/10.2807/1560-7917.ES.2015.20.39.30026.
2. Fontaine RE. *Describing epidemiologic data*. 2018. https://www.cdc.gov/eis/field-epi-manual/chapters/Describing-Epi-Data.html.

3. Soetens L, Hahné S, Wallinga J. Dot map cartograms for detection of infectious disease outbreaks: an application to Q fever, The Netherlands and pertussis, Germany. *Euro Surveill* 2017;**22**(26):30562. https://doi.org/10.2807/1560-7917.ES.2017.22.26.30562.
4. Boulos MNK. Descriptive review of geographic mapping of severe acute respiratory syndrome (SARS) on the Internet. *Int J Health Geogr* 2004;**3**(2). https://doi.org/10.1186/1476-072X-3-2.
5. Stone ER, Gabard AR, Groves AE, Lipkus IM. Effects of numerical versus foreground-only icon displays on understanding of risk magnitudes. *J Health Commun* 2015;**20**(10):1230–41. https://doi.org/10.1080/10810730.2015.1018594.
6. Wright AJ, Whitwell SCL, Takeichi C, Hankins M, Marteau TM. The impact of numeracy on reactions to different graphic risk presentation formats: an experimental analogue study. *Br J Health Psychol* 2009;**14**(1):107–25. https://doi.org/10.1348/135910708X304432.
7. MacEachren AM. The role of complexity and symbolization method in thematic map effectiveness. *Ann Assoc Am Geogr* 1982;**72**(4):495–513. https://doi.org/10.1111/j.1467-8306.1982.tb01841.x.
8. Baddeley A, Eysenck MW, Anderson MC. *Memory*. London: Routledge; 2020. https://doi.org/10.4324/9780429449642.
9. Rowland CA. The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychol Bull* 2014;**140**(6):1432–63. https://doi.org/10.1037/a0037559.
10. Reyna VF. A theory of medical decision making and health: fuzzy trace theory. *Med Decis Making* 2008;**28**(6):850–65. https://doi.org/10.1177/0272989X08327066.
11. Weinstein ND. What does it mean to understand a risk? Evaluating risk comprehension. *J Natl Cancer Inst Monogr* 1999;**25**:15–20. https://doi.org/10.1093/oxfordjournals.jncimonographs.a024192.
12. Vessey I. Cognitive fit: a theory-based analysis of the graphs versus tables literature. *Decis Sci J* 1991;**22**(2):219–40. https://doi.org/10.1111/J.1540-5915.1991.TB00344.X.
13. Lipkus IM. Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. *Med Decis Making* 2007;**27**(5):696–713. https://doi.org/10.1177/0272989X07307271.
14. Ozimec A-M, Natter M, Reutterer T. Geographical information systems-based marketing decisions: effects of alternative visualizations on decision quality. *J Market* 2010;**74**:94–110. http://www.jstor.org/stable/25764285.
15. Few S. Data visualization for human perception. In: Soegaard M, Dam RF, editors. *The encyclopedia of human-computer interaction*. 2nd ed. Arahus, Denmark: The Interaction Design Foundation; 2012.
16. Nuamah JK, Seong Y, Jiang S, Park E, Mountjoy D. Evaluating effectiveness of information visualizations using cognitive fit theory: a neuroergonomics approach. *Appl Ergon* 2020;**88**:103173. https://doi.org/10.1016/j.apergo.2020.103173.
17. Umanath NS, Scamell RW. An experimental evaluation of the impact of data display format on recall performance. *Commun ACM* 1988;**31**(5):562–70. https://doi.org/10.1145/42411.42419.
18. Mennecke BE, Crossland MD, Killingsworth BL. Is a map more than a picture? The role of SDSS technology, subject characteristics, and problem complexity on map reading and problem solving. *MIS Q– Manag Inf Syst* 2000;**24**(4):601–29. https://doi.org/10.2307/3250949.
19. Tversky A, Kahneman D. Availability: a heuristic for judging frequency and probability. *Cognit Psychol* 1973;**5**(2):207–32. https://doi.org/10.1016/0010-0285(73)90033-9.
20. Chua HF, Yates JF, Shah P. Risk avoidance: graphs versus numbers. *Mem Cognit* 2006;**34**(2):399–410. https://doi.org/10.3758/BF03193417.
21. Katapodi MC, Dodd MJ, Facione NC, Humphreys JC, Lee KA. Why some women have an optimistic or a pessimistic bias about their breast cancer risk: experiences, heuristics, and knowledge of risk factors. *Cancer Nurs* 2010;**33**(1):64–73. https://doi.org/10.1097/NCC.0b013e3181b430f9.
22. Shepperd JA, Lipsey NP, Pachur T, Waters EA. Understanding the cognitive and affective mechanisms that underlie proxy risk perceptions among caregivers of asthmatic children. *Med Decis Making* 2018;**38**(5):562–72. https://doi.org/10.1177/0272989X18759933.
23. Centers for Disease Control and Prevention. *Zika virus: statistics and maps*. www.cdc.gov/zika/reporting/.
24. Amazon Web Service. *Mechanical Turk concepts*. https://docs.aws.amazon.com/AWSMechTurk/latest/RequesterUI/mechanical-turk-concepts.html.
25. Juergens C. Trustworthy COVID-19 mapping: geo-spatial data literacy aspects of choropleth maps. *J Cartogr Geogr Inf* 2020;**70**:155–61. https://doi.org/10.1007/s42489-020-00057-w.
26. Tyner J. *Introduction to thematic cartography*. Englewood Cliffs, NY: Prentice-Hall; 1992.
27. Polland J. *A detailed look at how Americans travel within the US. Business Insider*. http://www.businessinsider.com/the-mostpopular-us-states-for-tourism-2014-10.
28. Yang ZJ, Richard LN, Harrison TM, Seo M. Applying the risk information seeking and processing model to examine support for climate change mitigation policy. *Sci Commun* 2014;**36**(3):296–324. https://doi.org/10.1177/1075547014525350.
29. Diamantopoulos A, Sarstedt M, Fuchs C, Wilczynski P, Kaiser S. Guidelines for choosing between multi-item and single-item scales for construct measurement: a predictive validity perspective. *J Acad Market Sci* 2012;**40**:434–49. https://doi.org/10.1007/s11747-011-0300-3.

Public Health 202 (2022) 66–73

Contents lists available at ScienceDirect

Public Health

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

Original Research

Redefining avoidable and inappropriate admissions

B. Clubbs Coldron [a, *], S. MacRury [a, b], V. Coates [c, d], A. Khamis [e, f]

[a] University of Highlands and Islands, Division of Rural Health and Wellbeing, Centre for Health Science, Old Perth Road, Inverness, IV2 3JH, UK
[b] Raigmore Hospital, Old Perth Road, Inverness, IV2 3UJ, UK
[c] School of Nursing, Ulster University, Northland Road, Derry, BT48 7JL, UK
[d] Western Health and Social Care Trust, Altnagelvin Area Hospital, Glenshane Road Londonderry, BT47 6SB, UK
[e] Letterkenny University Hospital, Kilmacrennan Rd, Ballyboe Glencar, Letterkenny, Co. Donegal, F92 AE81, Ireland
[f] National University of Ireland, Galway, University Road, Galway, H91 CF50, Ireland

ARTICLE INFO

Article history:
Received 21 April 2021
Received in revised form
18 October 2021
Accepted 4 November 2021
Available online 11 December 2021

Keywords:
Avoidable admissions
Inappropriate admissions
Chronic illness
Emergency care
Health policy
Health policy and planning
Participation
Policy process
Prevention
Diabetes

ABSTRACT

Objectives: Focusing on policy discourse in the United Kingdom, we examine the chain of causation that is characteristic of the ways in which the concepts of avoidability and inappropriateness are defined and used in these contexts. With a particular focus on diabetes complications, we aim to elucidate the way in which avoidable admission to hospital is conceptualised, measured, and applied to policy development and implementation and build a more inclusive model of identification as a basis for further research in this area.

Study design: Discourse analysis was used in combination with a scoping review.

Methods: We searched the online databases of the UK Houses of Parliament Hansard, Official reports of the Northern Ireland Assembly and transcripts of the Scottish Parliament in October 2021. We also conducted an electronic search in October 2021 on MEDLINE, PubMed, Google Scholar, EMBASE, CINAHL and The Cochrane Library to review the available literature. In addition, an analysis of policies in place in Scotland, England and Northern Ireland relating to urgent diabetes care was conducted.

Results: 'Avoidable' and 'inappropriate' hospital admissions are categories used in health policy and practice internationally as ways of identifying targets for interventions intending to reduce the burden of care. Diabetes mellitus is a chronic condition that is often seen as a costly and avoidable use of health care services and so is a frequent target of such policies.

Avoidable admission is interpreted as having a very long chain of causation. The assumption is that people requiring unscheduled hospital admission could have taken steps to prevent the onset of diabetes, or associated complications, arising in the first place. Definitions focus on primary and secondary prevention and largely place responsibility on the individual and their behaviour rather than on structural or social factors. Inadequate or inappropriate care prehospital or in the emergency department is seldom considered as a potential cause of avoidable admissions. Procedural definitions of avoidable admission are proposed whereby health care professionals and people living with diabetes collaborate to identify avoidable admissions in clinical audit rather than using statistical rates of avoidable admission within isolation in policy development and implementation.

Conclusions: Avoidability and inappropriateness are characteristics of cases in which conduct of the individual or attendant health care professionals was a proximate cause of hospital admission, and but for such conduct, admission could have been avoided. This process of definition seeks to provide a basis for contextualised and considered evaluation of where there are problems in care and where there are reasonable opportunities for prevention.

* Corresponding author. Tel.: +07507913846.
E-mail address: Benjamin.ClubbsColdron@uhi.ac.uk (B. Clubbs Coldron).

https://doi.org/10.1016/j.puhe.2021.11.004

Introduction

Diabetes Mellitus, and in particular type 2 diabetes (T2D), is an international public health issue. It presents a significant challenge to governments, clinicians and individuals self-managing the condition. It is estimated that worldwide roughly 463 million adults were living with diabetes in 2019 and that by 2045 this will rise to 700 million, according to the International Diabetes Federation.[1]

Reducing so-called 'avoidable' and 'inappropriate' admissions is a priority for the NHS (National Health Service) in the United Kingdom, as well as for health care systems across the globe.[2] Diabetes-related unscheduled care admissions are frequently assumed to be avoidable and/or inappropriate in statistical analysis.[3] Better self-management, empowered patients and increasingly sophisticated and personalised clinical decision making in theory has the potential to reduce demand and expenditure on services by ensuring avoidable conditions are prevented.[4,5] Public health initiatives to improve diet and exercise routines, education, reductions in socioeconomic disadvantage, integration, and increased accessibility of services (primary prevention measures) can reduce the risk of Type 2 diabetes (T2D) developing in the first place.[6] Intensive lifestyle modification and improved self-management (secondary prevention measures) can also reduce the risk of developing diabetes complications for people with both Type 1 diabetes (T1D) and T2D.[7]

Public health initiatives to prevent the onset and complications of diabetes vary internationally but generally focus on prevention through lifestyle change. Across Europe, the U.S., China, Australia, Japan and India, people considered high risk of developing T2D are targeted with education on diet and exercise, and this has been shown to significantly reduce progression to T2D in a wide variety of settings.[8–14] In the USA, Canada, Chile, the UK and New Zealand individuals in lower-risk tiers are targeted via risk counselling and whole population strategies (e.g. socioeconomic policies aiming to reduce poverty, healthy food promotion and environmental/systems changes).[15–22] Population-level policies, systems, and environmental approaches, along with lifestyle intervention for those at high risk, are likely to be the best way to achieve the greatest level of impact.[23]

Avoidability and inappropriateness are not synonymous, and the distinction between the two is important as it provides a basis for identifying, targeting, and designing interventions and policy initiatives to reduce admission for people with diabetes. In addition, interventions must be evidence-based and carefully tailored for specific contexts in order to be effective.[24,25] Evidence about what is effective in reducing avoidable admissions is mixed and inconclusive, and misconceptions about what and what is not avoidable can lead to naïve or unrealistic expectations of what might be achieved.[25]

Based on the political discourse around diabetes policy in the United Kingdom, a variety of problematic assumptions appear to underly initiatives aiming to reduce avoidable admissions. For instance, policy makers aiming to reduce hospital admissions appear to assume that there is an optimum level of admission or referral to hospital and that fewer admissions or referrals indicates an improvement in health care delivery and efficiency.[25] Only a small number of primary care trusts (PCTs) in England have successfully reduced overall unscheduled hospital admissions despite numerous initiatives.[26] The rate of avoidable admission also varies considerably across different studies.[25,27]

For example, the 2013 Urgent and Emergency Care Review for England stated that '40% could have been helped just as well closer to home'[28] (p. 19). By contrast, the NHS England Next Steps on the Five-Year Forward View suggests that 'between 1.36% and 2.73% of people presenting at A&E could be diverted away from hospital'[29] (p. 14). NHS Digital suggest that '16.1% of ED attendances occurring between April 2015 and December 2017 were 'non-urgent'.[30] The Nuffield Trust identify Ambulatory care sensitive conditions and urgent care sensitive conditions combined as comprising 3.38% of admissions between 2019 and 2020.[31] This variation may be a result of random fluctuations, contextual factors and different ways of conceptualising and measuring avoidability.[32,33] This means evaluation and comparison are problematic when based on crude statistical rates of avoidable admission. Clinical audit, with extensive input from service users, may bring meaning to crude rates of avoidable admission and allow interventions to be better targeted.[9]

Methods

A discourse analysis was conducted on political debates in the UK Houses of Parliament, The Scottish Parliament and the Northern Ireland Assembly. We searched UK Parliamentary Scottish Parliament and Northern Ireland Assembly Hansard from 2000 to 2021. Keywords included 'Avoidable Admission'; 'Inappropriate admission'; 'Hospital' 'unscheduled admission' and 'Diabetes'. We only excluded extracts that were deemed after an initial review to mention these topics incidentally rather than it being a central topic of discussion. This was based on the principles of discourse analysis outlined by Lupton.[34] A total of 137 extracts from these debates were subjected to discourse analysis and organised into themes. Extracts were included on the basis that they related to reducing avoidable, preventable or inappropriate hospital admissions in diabetes-related cases.

Policy analysis was conducted on the most recent diabetes-related policy documents from the United Kingdom and Scotland, also based on the above inclusion criteria. In addition, we undertook a scoping review to determine the depth of the literature around definitions and usage of avoidability and inappropriateness in relation to hospital admission to provide a detailed overview of the ways that these terms are used, to whom they are applied, where and when, as well as the potential benefits and limitations in terms of developing effective policy.[35]

An electronic search was conducted in October 2021 on MEDLINE, PubMed, Google Scholar, EMBASE, CINAHL and The Cochrane Library to review the available literature. We searched the titles and abstracts of papers, and the time period covered was from 2000 to 2021. Thirty articles from across the globe that examined the discourse of avoidability, measurement of avoidable admissions, policy development and implementation were included (see Fig. 1). Keywords included 'Avoidable Admission'; 'Inappropriate admission'; 'Hospital' 'Unscheduled admission' and 'Diabetes'. This research did not require ethical approval.

Results

Defining avoidable and inappropriate admissions

The verb 'avoid', means to escape, evade, prevent, or obviate. Historically, the term 'avoidable' has been used in a pejorative sense. Historically the term was often used to attribute causality and blame when someone failed to avoid a given outcome.[36] An associated concept is 'inappropriateness', which refers to something that is unsuitable to the particular case; 'unfitting, improper'.[37] It has been used to distinguish between desirable or undesirable behaviours based on the pragmatics of a given context. Although the concepts of inappropriateness and avoidability have some overlap and are frequently used interchangeably in the discourse on diabetes-related hospital admissions, there are some important differences that require explanation.

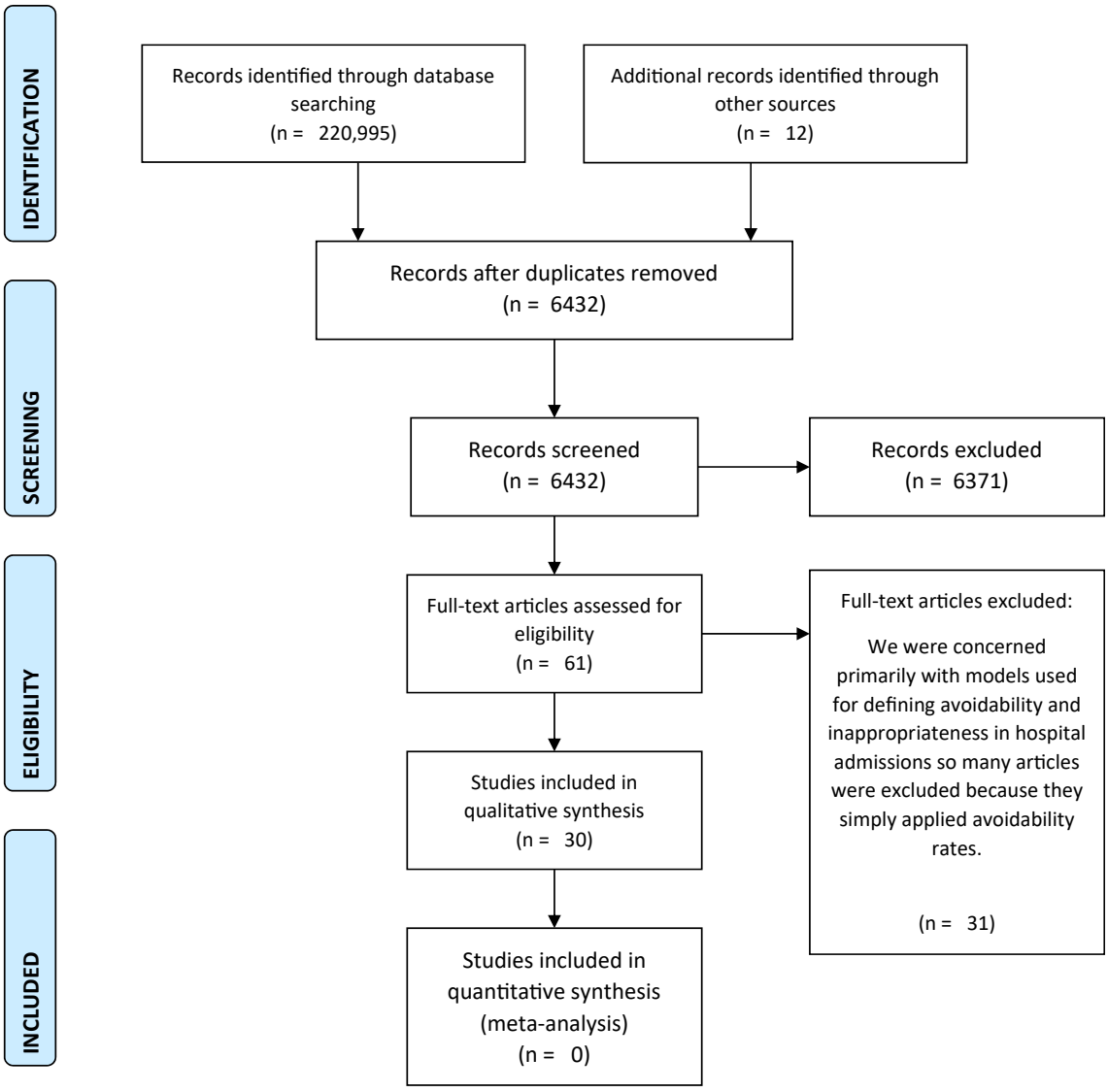


Fig. 1. Defining Avoidable and Inappropriate Hospital Admissions: A scoping review – PRISMA Flow diagram.

A case of diabetes-related hospital admission may, in theory, be avoidable in the sense that events in the chain of causation leading to hospital admission may have been preventable. With hindsight, it is almost always possible to identify an act or omission that contributed to hospitalisation, and it could have been avoided. In law, this is called a *novus actus intervenes*, or an intervening act that breaks the chain. However, the mere possibility of an intervening act (e.g. in public health, education, primary care, prehospital or emergency departments) does not necessarily render the eventual hospitalisation inappropriate. In the presenting circumstances of a medical emergency, it is, of course, appropriate and necessary that lifesaving care in the hospital setting is provided. Appropriateness is thus a contextualised concept that depends on the way decisions are made by HCPs in the agony of the moment. It is an evaluation of whether decisions leading to hospitalisation were reasonable. Conversely, avoidability merely depends on the possibility of breaking the chain of causation and does not necessarily ask whether avoidance was feasible in situ (Fig. 2).

Although an avoidable admission is not necessarily inappropriate, an inappropriate admission may, by definition, be avoidable. If hospitalisation is not required – perhaps because suitable treatment can be provided in the community or primary care – then the admission could be prevented by diversion to the appropriate services.[24]

Objectivised measures of appropriateness and avoidability

Objectivised methods of defining avoidable admissions come in two forms: checklist definitions based on a set of standardised criteria and definitions based on professional opinion and/or expert panels. Checklist models were initially developed in the USA to decide the hospital admissions that were appropriate for insurers to fund. An initial function of the language of appropriateness was to distinguish between deserving from undeserving service users. Checklist models are standardised lists of criteria, which indicate a set of symptoms or circumstances that necessitate hospital admission. Criteria usually relate to the severity of an individual's condition and the type and intensity of services provided.[24]

Prevalent checklist models are the AEP (Appropriateness Evaluation Protocol), the ISD-A (Intensity-Severity-Discharge Review System with Adult Criteria) and MCAP (Managed Care Appropriateness Protocol. A term used frequently in UK policy and academic discourse is Ambulatory Care Sensitive Conditions (ACSCs), denoting conditions that can be treated effectively in primary care.

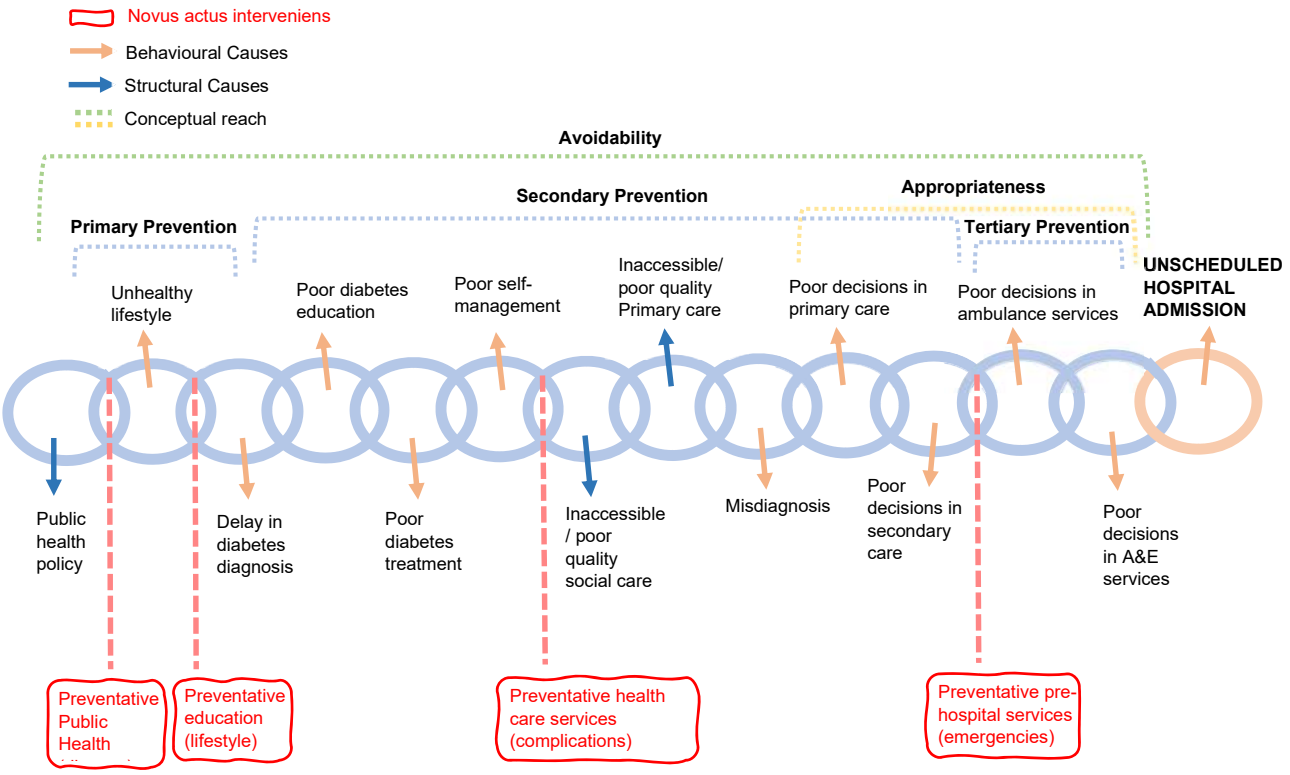


Fig. 2. The chain of causation in unscheduled diabetes-related hospital admissions.

Admissions for ACSCs are often assumed in statistical models to refer to avoidable admissions.[2] Checklist models produce easily quantifiable results and can help public health professionals and policy makers in evaluating and comparing services. The Appropriateness Evaluation Protocol (AEP) proposes admission criteria based only on physiological and laboratory parameters. However, the AEP has been found to be a poor predictor of mortality in all age groups. Researchers have advised that it not be used to evaluate the appropriateness of admissions.[38]

Another model used in the USA is the NYU algorithm[40] that assigns the probability that an ICD-9 diagnosis code associated with an Emergency Department visit falls into one of four categories:

1) a non-emergency (NE);
2) an emergency (defined as a problem requiring contact with the medical system within 12 h) treatable in an office visit (primary care treatable (PCT));
3) an emergency not treatable in an office visit but preventable or avoidable (EPA) and
4) an emergency that is not preventable or avoidable (ENPA).

The NYU algorithm excludes uncommon diagnoses and treats mental health and substance abuse diagnoses separately.[39] This model has been independently validated using hospitalisations and deaths as outcome measures, so it appears relatively robust. Ballard et al. found that because the NYU algorithm utilizes existing clinical data rather than time-intensive chart review, it can be easily applied in different times and settings at relatively low cost.[39] A similar model could be useful in evaluating and comparing avoidable admission rates between hospitals in the UK context rather than relying on ACSCs. However, the NYU Emergency department visit classifier, by the developer's own admission, is not appropriate for determining avoidable and inappropriate admissions in a way that might suggest appropriate changes in social, or primary and secondary health care delivery at the local or individual level.[39]

Checklist criteria combined with the use of algorithms are useful in constructing statistical estimates of the number of preventable hospital admissions in the UK and other jurisdictions for comparison and evaluation of performance. Diabetes complications are a prime example of some of the weaknesses of this approach. In reports produced by the Nuffield Trust,[3] avoidable admission rates for people with diabetes in the UK are built from the sum of three indicators: admissions for short-term diabetes complications; admissions for long-term diabetes complications; admissions for uncontrolled diabetes without complications.

The rate of avoidable admission for people with diabetes is defined as the number of hospital admissions with a primary diagnosis of diabetes, among people aged 15 years and over, per 100,000 population. The Information Service Division (ISD) of NHS Scotland, in presenting statistical analysis, similarly defines practically all hospital admissions for people with diabetes as avoidable.[40] The ISD explain that this is because routine monitoring, dietary modification and regular exercise can reduce the need for hospitalisation.[40] This methodology is based a problematic assumption that personal choices are the main driver of avoidable admissions. A more detailed classification system that does n' rely on lumping all diabetes complications into the category of avoidable, for example, based upon the NYU Categories, would perhaps be more useful (See Fig. 3).

Political discourse

The discourse that frames the classification of all diabetes-related hospital admissions as avoidable in the UK generally focuses on the ways in which risk of diabetes-related complications could be reduced by the person themselves through individual behavioural change. This places almost all responsibility for hospitalisation on the individual and does little to inform for positive change in clinical practise or health care delivery at the local and individual level. For this kind of change, more detailed auditing of

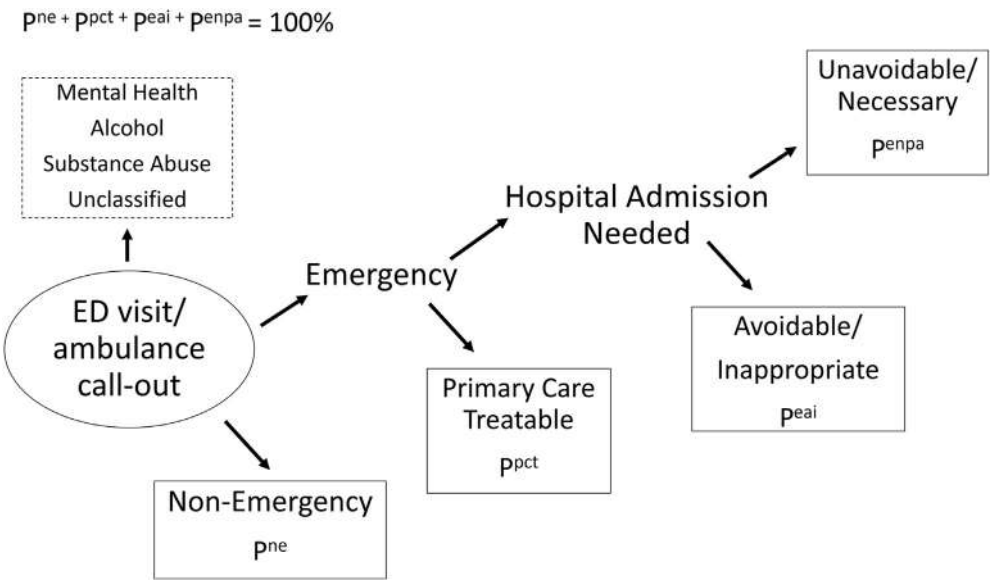


Fig. 3. NYU Algorithm adapted for use in the NHS System.

individual and local cases appears necessary alongside something like the NYU Algorithm to identify drivers of avoidable admissions at a more granular level and to improve care for particular chronic conditions such as diabetes.

The fact that the risk of diabetes complications can be reduced by improvements in lifestyle, including adherence to therapy and diet, is seen as sufficient to define all diabetes-related admissions as avoidable. This is a blanket generalisation that may not reflect the experience of people with diabetes. In some circumstances and for some individuals, it is not reasonable to expect them to change their way of life and the environments in which they live may not be exercise friendly or they may find it difficult to access healthy and nutritious food. The expectation that a person diagnosed with diabetes should increase their daily exercise, adapt their diet and frequently monitor and manage their glucose levels is often unrealistic in the absence of public health initiatives and involves assumptions about costs and benefits that may not take sufficient account of the individual's conception of the good life. In addition, reducing health risks does not necessarily provide a guarantee of a life free of diabetes-related complications.

This discourse shifts focus away from public health, primary and secondary care, and accessibility towards individual behaviour. These assumptions are embedded at the root level of the data and statistics and so are objectivised and become incontestable at the stage of political debate and policy development. We can contrast the way diabetes-related and smoking-related admissions are treated. Smoking-related conditions are not included en masse in the ISD avoidable admissions rate (COPD) even though smoking cessation can significantly reduce the risk. The logic of personal responsibility is inconsistently applied, indicating a special contempt for people with T2D in particular.

Using a statistical rate of avoidable admissions, based on checklist criteria, as an objective comparator between health providers is dubious because data recording and coding practices vary considerably within the UK. For example, coding of diabetes as a principal diagnosis versus a secondary diagnosis varies between Northern Ireland, England and Scotland, making direct comparisons difficult. Even within England, the comparison between regions is questionable.[26] The north of England, for instance, has higher avoidable admission rates even adjusted for deprivation (IMD); therefore, it is possible the variation is due to other factors such as disparities in primary, community and secondary care provision, health service accessibility or the wider determinants of health not included in the IMD.

Reported rates of avoidable admission, based on current definitions, may be difficult to compare between trusts or hospitals as institutions recording of cases involving diabetes complications are not necessarily recorded as such. Neither do such measures take account of differences in disease prevalence, local services, or culture.[11,26] The ISD-A and NYU algorithm, for example, do not consider the fact that there may be no other option in the local area for the individual except hospital.[19] In addition, the AEP is often amended and adjusted in practice, which means that assessments are unreliable.[26,41–43]

This is not to say that statistical methods of measuring avoidable and inappropriate admissions should not be used. It is that they are not very effective tools in evaluating local services or clinical practice in constructive ways. It is suggested that more detailed, local and individual audits of clinical practice in areas such as diabetes care could be a vital addition to existing measures of avoidable admissions. User-led definition and audit of avoidable cases may offer invaluable insights, and this should also factor into policy and practice looking to reduce avoidable admissions.[24]

In applying definitions of avoidability, professionals exercise subjective judgement as to whether a given admission is appropriate or not. Checklists and algorithms appear to be helpful as a rough outline of relevant considerations; however, they also objectify the outcomes of such judgements and can even embed prejudice about people with conditions such as diabetes. Appropriateness of a given admission depends on when the AEP or ISD-A are applied to each person's case and by whom.[24,44] The NYU method, although appearing to be more objective by the application of computerised algorithms, uses only clinical data, and this assumes that a wealth of information about individual and local circumstances from a service user perspective is irrelevant.

The method of defining avoidability through assessment by algorithms, experts and professionals is an exercise of judgement based on disciplinary knowledge, opinion, and experience. Classification may also be conducted by trained researchers. In both

procedures, the expertise and experience of people with diabetes themselves are not required.[11] This indicates that direct experience of diabetes is not adequately recognised as a useful or legitimate contribution to the process of definition.

Only two previous studies have included service users in assessing the avoidability and appropriateness of hospital admission.[24,32] These studies indicate that the term inappropriate, as applied to admissions, carries several negative connotations that service users are reluctant to apply.

For people with T2D, improved diet, lifestyle, and glucose control, usually at an early stage, can be associated with remission and a significant reduction the risk to health and wellbeing.[45] However, such measures cannot eliminate the risk of complications arising nor render diabetes-related hospital admission an impossibility. Therefore, a proportion of diabetes-related admissions must be necessary and appropriate. This means that statistics on avoidable admissions that include all diabetes-related hospital admissions are misleading without greater input of service users on what leads them through the hospital doors. Despite this, the policy discourse continues to place responsibility for diabetes-related hospital admission on people with diabetes themselves. This can frame the problem of overstretched services as within the power of individuals to change and beyond the scope of governmental responsibility. This has implications for both policy and the way in which admissions are viewed and dealt with by HCPs. It could also influence behaviour in the uptake of services.

Constructing the use of unscheduled care services by people with diabetes as problematic and 'inappropriate' in almost all cases could paradoxically create a culture in which patients are discouraged from seeking help when they need it to avoid judgement, resulting in more serious and preventable complications reaching crisis point. Greater efficiency (shorter stays in hospital for each admitted patient) may lead to more 'inefficiency' (greater number of avoidable emergency admissions).[24]

The expertise and experience of people using services due to diabetic complications are crucial to understanding the context in which unscheduled admissions occur and developing appropriate reforms to reduce avoidable admissions however they are defined.[46] This is particularly important at a time when law and policy emphasize a commitment to person-centred care and 'nothing about me without me'.[47] More inclusive and democratic epistemologies expand the range of evidence and data on which our knowledge draws and can thus provide a more holistic, long-term view of the factors that contribute to hospital admissions. The inclusion of perspectives of people with direct experience will help to build a more comprehensive picture of how best to respond to the interests and perspectives of people with diabetes where prevention is possible, reasonable and desirable in light of competing priorities.[11]

Discussion

There is limited evidence on the effectiveness of interventions aimed at reducing unplanned admissions in diabetes cases to date. Interventions have generally been focused on different stages along the patient journey, from preventive management of people at high risk of admission, through to services that manage acute diabetic complications without resorting to hospital admission.[48] Interventions often focus on individual patients and seek to develop capabilities for self-management.[49] This reflects a long view of causation of admission, with the privilege of hindsight and without an intimate understanding of surrounding circumstances of the person.

For researchers, HCPs and policymakers, avoidable admissions are usually based on the idea that people with diabetes could have taken steps to prevent the disease and associated complications arising in the first place.[50] It is helpful to visualise this as a chain of causation that is cut in different ways according to diverse perspectives (see Fig. 1).

Defining avoidability in diabetes-related cases and persons with other chronic health conditions is a judgement call on which events are relevant in the chain of causation leading to admission (see Fig. 1). An exclusively medical perspective can mean that a person's unique circumstances and context are inadequately considered, and thus relevant knowledge is excluded from the decision-making process. A procedural definition of avoidability for use in clinical audit that incorporates both HCP and service user perspectives is proposed as an alternative to the statistical comparison of rates of avoidable admissions. The criterion for avoidable admissions is based on a test for causation in medical negligence cases. It is a procedure for reaching a considered judgement and not an objective test.

To identify a case of hospitalisation as avoidable or inappropriate, it should pass a version of the 'but for' and Bolam/Bolitho tests in law.[51,52] This allows a considered approach to the question of causation and provides service users with an opportunity to contribute to the narrative of their admissions:

1. Avoidable hospital admissions can broadly be identified where:
   a. Admission would not have occurred if the policymakers, HCPs and/or service users had taken all reasonable steps to ensure prevention, diagnosis, and management in the community and;
   b. Unscheduled hospital admission directly followed (i.e. it is a proximate cause of admission rather than far removed down the chain of causation – see Fig. 1).
2. Inappropriate admissions may be identified where
   a. Decisions were made in social, primary, secondary, pre-hospital and/or emergency services that other reasonable clinicians would regard as unsuitable to the patient's needs and circumstance and;
   b. Unscheduled hospital admission directly followed.

The test may, in practice, still be vulnerable to the tendency to place undue emphasis on HCP perspectives in determining what 'skilled' or 'preventative' practices are. Frequently the voice of clinical expertise can become dominant. If the knowledge of those with direct experience could be accorded equal value as clinical expertise, and if both HCPs and service uses are equally represented in the retrospective classification of cases, then the interests and perspectives of all relevant stakeholders can be given due consideration. This should be a measure that is undertaken in addition to the use of something like the NYU algorithm and could be used to define some of the key indicators such as 'non-emergency' 'primary care treatable' and 'avoidable/inappropriate' (see Fig. 3). An issue with the methods proposed is that they could produce outcomes that are not amenable to straightforward statistical analysis. They could also be viewed as inefficient. However, they may prove their worth in coconstructing best practices that can improve the avoidance of hospital admission in local contexts and may be useful in providing more detailed comparisons between different approaches. In this way, it may be very useful in both developing policies and clinical commissioning, as well as tailoring policy to localities and communities.

Conclusions

In policy and practice, avoidability and inappropriateness have become so familiar that they are not seen as a way of 'ordering', but as an 'order inherent in the phenomena'.[26] Here, we have

attempted to set aside the taken-for-granted definitions of these concepts and examine alternative and, we suggest, potentially more productive practices of categorisation.

Objective comparisons between services based on avoidable or inappropriate admissions remains an ideal rather than a reality. Rather, avoidability in the policy discourse appears to be used to frame the treatment of people with chronic conditions in hospitals as an unnecessary distribution of resources.

Although there may have been a degree of selection bias in the materials analysed, it appears that in a wide variety of political discourses, the focus on diabetes, as a particular contributor to avoidable admission rates. This reflects a tacit assumption that diabetes complications are primarily caused by individual lifestyle choices. This can function to exclude the legitimate knowledge of people with both T1D and T2D and justify the rationing of services. This is highlighted by a more beneficent approach to policy and practice vis-à-vis people, and in particular, children with T1D who tend to be viewed more as random victims of a disease. Even in these cases, a lack of glycaemic control is often seen as a culpable reason for unnecessary admissions.

If the use of unscheduled care services by people with diabetes is unilaterally defined as 'problematic and inappropriate' in almost all cases, we could discourage people from seeking help when they need it, resulting in more serious and preventable complications reaching crisis point. Ensuring that people seek help and that there is the capacity to deal with early intervention is crucial. Timely access to and the response of services is important not only in diabetes but other conditions too. Using a model similar to the NYU algorithm to identify problematic hospital admission rates in combination with a more intensive auditing procedure based on collaborative definitions of avoidability and inappropriateness as modelled above could be more useful in improving service provision and preventing avoidable hospitalisation in chronic conditions such as diabetes.

Author statements

Acknowledgements

The views and opinions expressed in this report do not necessarily reflect those of the European Commission or the Special EU Programmes Body (SEUPB).

Ethical approval

No ethical approval was required as the research consisted of documentary analysis and literature review.

Funding

This work is part of a project funded by the INTERREG VA Programme, managed by the Special EU Programmes Body (SEUPB).

Competing interest

There are no conflicts of interest to declare.

References

1. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, Colagiuri S, Guariguata L, Motala AA, Ogurtsova K, Shaw JE, Bright D, Williams R, IDF Diabetes Atlas Committee. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: results from the international diabetes federation diabetes atlas, 9th edition. *Diabetes Res Clin Pract* 2019;**157**:107843. Nov.
2. Purdy S, Griffin T, Salisbury C, et al. Ambulatory care sensitive conditions: terminology and disease coding need to be more specific to aid policy makers and clinicians. *Publ Health* 2010;**123**:169–73. https://doi.org/10.1016/j.puhe.2008.11.001. 2009.
3. NHS Digital. *Ambulatory care sensitive conditions (ACSC)*. 2021. Available at: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/innovative-uses-of-data/demand-on-healthcare/ambulatory-care-sensitive-conditions#seasonality-and-trends. [Accessed 11 October 2021].
4. Goodwin N, Sonola L, Thiel V, Kodner D. *Co-ordinated care for people with complex chronic conditions: key lessons and markers for success*. London: The King's Fund; 2013. Available at: www.kingsfund.org.uk/publications/co-ordinated-care-people-complex-chronic-conditions. [Accessed 21 January 2014].
5. Goodwin N, Smith J, Davies A, Perry C, Rosen R, Dixon A, Dixon J, Ham C. *A report to the Department of Health and the NHS Future Forum. Integrated care for patients and populations: improving outcomes by working together*. London: The King's Fund and Nuffield Trust; 2012. Available at: www.kingsfund.org.uk/sites/files/kf/integrated-care-patientspopulationspaper-nuffield-trust-kings-fund-january-2012.pdf. [Accessed 15 January 2014].
6. Spencer Bonilla G, Rodriguez-Gutierrez R, Montori VM. What we don't talk about when we talk about preventing type 2 diabetes—addressing socioeconomic disadvantage. *JAMA Intern Med* 2016;**176**(8):1053–4.
7. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002;**346**:393–403 [Web].
8. Thomas C, Sadler S, Breeze P, Squires H, Gillett M, Brennan A. Assessing the potential return on investment of the proposed UK NHS diabetes prevention programme in different population subgroups: an economic evaluation. *Br Med J Open.* 2017;**7**(8). https://doi.org/10.1136/bmjopen-2016-014953. e014953.
9. Makrilakis K, Liatis S, Grammatikou S, Perrea D, Katsilambros N. Implementation and effectiveness of the first community lifestyle intervention programme to prevent type 2 diabetes in Greece. The DE-PLAN study. *Diabet Med* 2010;**27**(4):459–65. https://doi.org/10.1111/j.1464-5491.2010.02918.x.
10. Diabetes Prevention Program Research Group. The 10-year costeffectiveness of lifestyle intervention or metformin for diabetes prevention: an intent-to-treat analysis of the DPP/DPPOS. *Diabetes Care* 2012;**35**(4):723–30. https://doi.org/10.2337/dc11-1468.
11. Gong Q, Zhang P, Wang J, Ma J, An Y, Chen Y, et al. Morbidity and mortality after lifestyle intervention for people with impaired glucose tolerance: 30-year results of the Da Qing Diabetes Prevention Outcome Study. *Lancet Diabetes Endocrinol* 2019 june;**7**(6):452–61. https://doi.org/10.1016/S2213-8587(19)30093-2. Epud 2019 Apr 26. PMID: 31036503; PMCID: PMC8172050.
12. Laatikainen T, Dunbar JA, Chapman A, Kilkkinen A, Vartiainen E, Heistaro S, et al. Prevention of type 2 diabetes by lifestyle intervention in an Australian primary health care setting: greater Green Triangle (GGT) Diabetes Prevention Project. *BMC Publ Health* 2007;**7**(249). https://doi.org/10.1186/1471-2458-7-249.
13. Kosaka K, Noda M, Kuzuya T. Prevention of type 2 diabetes by Lifestyle intervention: a Japanese trial in IGT males. *Diabetes Res Clin Pract* 2005;**67**(2):152–62. https://doi.org/10.1016/j.diabres.2004.06.010.
14. Ramachandran A, Snehalatha C, Mary S, Mukesh B, Bhaskar AD, Vijay V. The Indian Diabetes Prevention Programme shows that lifestyle modification and metformin prevent type 2 diabetes in Asian Indian subjects with impaired glucose tolerance (IDPP-1). *Diabetologia* 2006;**49**(2):289–97. https://doi.org/10.1007/s00125-005-0097-z.
15. Andreyeva T, Long MW, Brownell KD. The impact of food prices on consumption: a systematic review of research on the price elasticity of demand for food. *Am J Publ Health* 2010;**100**(2):216–22. https://doi.org/10.2105/ajph.2008.151415.
16. Calancie L, Leeman J, Jilcott Pitts SB, Khan LK, Fleischhacker S, Evenson KR, et al. Nutrition-related policy and environmental strategies to prevent obesity in rural communities: a systematic review of the literature, 2002-2013. *Prev Chronic Dis* 2015;**12**:E57. https://doi.org/10.5888/pcd12.140540.
17. Olstad DL, Teychenne M, Minaker LM, Taber DR, Raine KD, Nykiforuk CI, et al. Can policy ameliorate socioeconomic inequities in obesity and obesity-related behaviours? A systematic review of the impact of universal policies on adults and children. *Obes Rev* 2016;**17**(12):1198–217. https://doi.org/10.1111/obr.12457.
18. Lichtenstein AH, Carson JS, Johnson RK, Kris-Etherton PM, Pappas A, Rupp L, et al. Food-intake patterns assessed by using front-of-pack labeling program criteria associated with better diet quality and lower cardiometabolic risk. *Am J Clin Nutr* 2014;**99**(3):454–62. https://doi.org/10.3945/ajcn.113.071407.
19. Hyseni L, Atkinson M, Bromley H, Orton L, Lloyd-Williams F, McGill R, et al. The effects of policy actions to improve population dietary patterns and prevent diet-related non-communicable diseases: scoping review. *Eur J Clin Nutr* 2017;**71**(6):694–711. https://doi.org/10.1038/ejcn.2016.234.
20. Niebylski ML, Redburn KA, Duhaney T, Campbell NR. Healthy food subsidies and unhealthy food taxation: a systematic review of the evidence. *Nutrition* 2015;**31**(6):787–95. https://doi.org/10.1016/j.nut.2014.12.010.
21. Albright A, Devlin H, Zhang X. Integrating nutrition therapy into community-based diabetes prevention programs (pp 579. In: Franz MJ, Evert AB, editors. *American diabetes association guide to nutrition therapy for diabetes*. 3rd ed. 2012.
22. Mayne SL, Auchincloss AH, Michael YL. Impact of policy and built environment changes on obesity-related outcomes: a systematic review of naturally occurring experiments. *Obes Rev* 2015;**16**(5):362–75. https://doi.org/10.1111/obr.12269.

23. 2020 Gruss SM, Nhim K, Gregg E, Bell M, Luman E, Albright A. Public health approaches to type 2 diabetes prevention: the US national diabetes prevention program and beyond. *Curr Diabetes Rep* 2019 Aug 5;**19**(9):78. https://doi.org/10.1007/s11892-019-1200-z. Erratum in: Curr Diab Rep. Jun 27;20(8):36. PMID: 31385061; PMCID: PMC6682852.
24. Thwaites R, Glasby J, Mesurier N, Littlechild R. Room for one more? A review of the literature on 'inappropriate' admissions to hospital for older people in the English NHS. *Health Soc Care Community* 2017;**25**(1):1–10.
25. Roland M, Abel G. Reducing emergency admissions: are we on the right track? *BMJ* 2012;**345**:e6017.
26. Dr Foster Intelligence. *Reducing avoidable emergency admissions Analysis of the impact of ambulatory care sensitive conditions in England*. 2018. Available at: https://www.evidence.nhs.uk/search?ps=40&q=Emergency%20Admissions%20for%20Ambulatory%20Care%20Sensitive%20Conditions%20-characteristics%20and%20trends%20at%20national%20level.%20. [Accessed 11 October 2021].
27. Gillam S. 'Rising hospital admissions: can the tide be stemmed? (editorial)'. *Br Med J* 2010;**340**:c636.
28. NHS England. *High quality care for all, now and for future generations: transforming urgent and emergency care services in England - urgent and Emergency Care Review*. 2013. Available at: https://www.nhs.uk/NHSEngland/keogh-review/documents/UECR.Ph1Report.Appendix%201.EvBase.FV.pdf. [Accessed 11 October 2021].
29. NHS England. *Next steps on the NHS five year forward view*. 2017. Available at: https://www.england.nhs.uk/publication/next-steps-on-the-nhs-five-year-forward-view/.
30. NHS Digital. *Unnecessary A and E attendances*. NHS Digital; 2021. https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/innovative-uses-ofdata/demand-on-healthcare/unnecessary-a-and-e-attend. [Accessed 11 October 2021].
31. Nuffield Trust. *Potentially preventable emergency admissions*. 2021. Available at: https://www.nuffieldtrust.org.uk/resource/potentially-preventable-emergency-hospital-admissions. [Accessed 11 October 2021].
32. Glasby J, Littlechild R, Le Mesurier N, Thwaites R. Who knows best? Older people's and practitioner contributions to understanding and preventing avoidable hospital admissions. *Health Econ Pol Law* 2019:1–22.
33. Parkinson B, Meacock R, Checkland K, Sutton M. Clarifying the concept of avoidable emergency department attendance. *J Health Serv Res Pol* 2021;**26**(1):68–73.
34. Lupton Deborah. Discourse analysis: a new methodology for understanding the ideologies of health and illness. *Aust J Publ Health* 1992;**16**(2):145–50.
35. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med Res Methodol* 2018;**18**:143.
36. Avoidable. In: *Oxford online dictionary*; 2019. Retrieved from: https://www.oed.com/view/Entry/13797?redirectedFrom=avoidable&.
37. Inappropriate. In: *Oxford online dictionary*; 2019. Retrieved from: https://www.oed.com/view/Entry/93142?redirectedFrom=inappropriate&.
38. O'Regan NA, Healy L, O Cathail M, Law TW, O'Carroll G, Clare J, Timmons S, O'Connor KA. The Appropriateness Evaluation Protocol is a poor predictor of in-hospital mortality. *Ir J Med Sci* 2014;**183**(3):417–21. https://doi.org/10.1007/s11845-013-1031-x. Epub 2013 Oct 30. PMID: 24170692.
39. Ballard DW, Price M, Fung V, Brand R, Reed ME, Fireman B, Newhouse JP, Selby JV, Hsu J. Validation of an algorithm for categorizing the severity of hospital emergency department visits. *Med Care* 2010;**48**(1):58–63. https://doi.org/10.1097/MLR.0b013e3181bd49ad.
40. NHS Scotland, Information Service Division. *Avoidable admissions*. 2019. Retrieved from, https://www.isdscotland.org/health-topics/public-health/publications/2017-02-21/Avoidable-Admissions.asp.
41. Littlechild R, Glasby J. Emergency hospital admissions: older patients' perceptions. *Educ Ageing* 2001;**16**(1):77–89.
42. Houghton A, Bowling A, Jones I, Clarke K. Appropriateness of admission and the last 24 hours of hospital care in medical wards in an east London teaching group hospital. *Int J Qual Health Care* 1996;**8**(6):543–53.
43. Coast J, Inglis A, Morgan K, Gray S, Kammerling M, Frankel S. The hospital admissions study in England: are there alternatives to emergency hospital admission? *J Epidemiol Community Health* 1995;**49**:194–9.
44. Tsang P, Severs M. A study of appropriateness of acute geriatric admissions and an assessment of the Appropriateness Evaluation Protocol. *J Roy Coll Phys Lond* 1995;**29**(4):311–4.
45. Mccombie L, Leslie W, Taylor R, Kennon B, Sattar N, Lean M. Beating type 2 diabetes into remission. *BMJ* 2017;**358**:J4030.
46. Blunt I, Bardsley M, Dixon J. *Trends in emergency admissions in England 2004–2009: is greater efficiency breeding inefficiency?* London: Nuffield Trust; 2010.
47. NHS Scotland. *Person-centred care: guidance for non-executive directors*. Edinburgh: The Scottish Government; 2016. from: https://www.gov.scot/publications/person-centred-care-non-executive-directors/. [Accessed 10 November 2020].
48. Turner S. Studying organization through Levi Strauss's structuralism. In: Morgan G, editor. *Beyond method: 189–201*. Beverly Hills, CA: Sage; 1983.
49. *Action for diabetes*. NHS England; 2014. from: https://www.england.nhs.uk/rightcare/wp-content/uploads/sites/40/2016/08/act-for-diabetes-31-01.pdf. [Accessed 17 September 2020].
50. The Scottish Government. *Diabetes improvement plan, health and social care*. November. 2014. from: file:///C:/Users/EO03BC/Downloads/00464461%20(5).pdf. [Accessed 17 September 2020].
51. Todd NV. Medical negligence. An overview of legal theory and neurosurgical practice: causation. *Br J Neurosurg* 2014;**28**(3):315–9.
52. Panting G. Clinical negligence. *Orthop Traumatol* 2015;**29**(4):268–72.

Public Health 202 (2022) 100–105

Contents lists available at ScienceDirect

Public Health

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

Original Research

Validation of the Hamilton Depression Rating Scale (HDRS) in the Tunisian dialect

N. Cheffi [a], O. Chakroun-Walha [b, *], R. Sellami [a], R. Ouali [a], D. Mnif [a], F. Guermazi [a], F. Issaoui [b], M. Lajmi [b], B. Benamar [b], J. Damak [c], N. Rekik [b], J. Masmoudi [a]

[a] Department of Psychiatry A, University Hospital Hédi Chaker Sfax, Tunisia
[b] Emergency Department, University Hospital Habib Bourguiba Sfax, Tunisia
[c] Department of Community Health and Epidemiology, University Hospital Hedi Chaker Sfax, Tunisia

ARTICLE INFO

Article history:
Received 11 February 2021
Received in revised form
23 October 2021
Accepted 5 November 2021
Available online 20 December 2021

Keywords:
HDRS scale
Suicide
Validation
Depression
Tunisia

ABSTRACT

Objectives: The Hamilton Depression Rating Scale (HDRS) is one of the most frequently used depression assessment scales. In Tunisia, psychiatrists commonly use this scale in a Tunisian dialect. However, to the best of our knowledge, this scale has never been validated in Tunisia. This study aims to investigate the reliability and the validity of the HDRS among Tunisian patients who have been hospitalised for a suicide attempt. A secondary objective is to describe the sociodemographic characteristics of the study population.

Study design: This is a cross-sectional study performed in the emergency department.

Methods: Patients who were hospitalised for a suicide attempt were eligible for inclusion in this study. The Tunisian version of the HDRS was developed using a forward-backward translation procedure. Psychometric properties of the Tunisian version of the HDRS were tested, including (i) construct validity with a confirmatory one-factor analysis; (ii) internal validity with Pearson correlations and Cronbach alpha coefficients; and (iii) external validity by correlations with the Patient Health Quality-9 (PHQ-9) scale. We used the Receiver-Operating Characteristic (ROC) curve to analyse the correlation between the total HDRS score and the presence of depression according to the PHQ-9.

Results: In total, 101 participants were enrolled in this study. The principal component analysis (PCA) type factor analysis with varimax rotation found a high–grade correlation between HDRS individual items and the total score. The total variance, explained by five factors, was 64.4%. Cronbach's standardised alpha coefficient was 0.86 for the overall scale. Correlations between the total HDRS score and the PHQ-9 score, and its various items, were significant. The ROC curve analysis showed good sensitivity (80.8%) and specificity (91.1%).

Conclusion: The Tunisian version of the HDRS is an acceptable instrument to screen depression in individuals who have attempted suicide.

Introduction

According to the World Health Organisation (WHO), depressive disorders constitute one of the leading causes of disability.[1] Each year, depression affects approximately 5% of the population worldwide.[2] Suicide is a major public health problem and is the second leading cause of death in people aged 15–29 years.[3] In Tunisia, there has been an increase in suicides following the revolution in 2011.[4] Considering the associations between depression and suicide, the diagnosis and treatment of depression is especially important in suicide prevention programmes.[5,6]

The assessment of depression is part of the routine care in clinical practice. The instruments used for the assessment should be clinically meaningful, reliable, valid and sensitive to the target population.[7] The ideal instrument should, therefore, encompass all aspects of the psychometric assessment and be valid for all cultures, ages, sexes, socio-economic and linguistic backgrounds.[8] The instrument must be applicable in different contexts to allow comparisons between multiple groups within one country and across countries.[8]

* Corresponding author. Tel.: +216 58 12 76 93.
E-mail address: chakroun_olfa@medecinesfax.org (O. Chakroun-Walha).

https://doi.org/10.1016/j.puhe.2021.11.003
0033-3506/

The Hamilton Depression Rating Scale (HDRS), published in 1960, is one of the most frequently used depression assessment scales, helping to determine the severity of a depressive state, responses to treatments and also aiding diagnosis of depression.[9] Since its publication, it has become the gold standard for the assessment of depression.[10] However, the HDRS is not a diagnostic instrument.[11]

Many versions of the HDRS exist, including the original version, with 21 items[9] (Hamilton, 1960), and one with 17 items (Hamilton, 1967), which has been translated and validated in the french version consists of 3 factors,[12] Turkish[13] Lebanese[14] and Chinese.[15] The Lebanese version consists of four factors, explaining 58.8% of the total variance.[14] The Chinese version consists of five factors, accounting for 52.4% of the total variance.[15] The Turkish version consists of six factors, explaining 61.3% of the total variance.[13]

In Tunisia, psychiatrists commonly use the HDRS scale in a Tunisian dialect; however, this scale has never been appropriately validated in Tunisia. Therefore, this study aims to investigate the reliability and validity of the HDRS among Tunisian patients who were hospitalised for a suicide attempt.

Methods

Clinical setting

This cross-sectional study was performed between July and October 2019 in the emergency department (ED) of the university hospital Habib Bourguiba, in Sfax, Tunisia. Informed consent was obtained after explanations about the aim and methods of the study. Patients had the right to accept or refuse participation in the study.

Inclusion criteria

This study was carried out with patients who were hospitalised in the ED for a suicide attempt. The emergency physician evaluated the patient's vital signs, the need for gastric decontamination, the need for blood tests, antidotes or other required treatments. When surveillance in the ED was no longer necessary (i.e. clinical stabilisation), the patient's eligibility to participate in this study was assessed. Patients with a life-threatening condition, those aged <18 years and those with difficulty communicating were not included in the study. Patients with a previous history of schizophrenia were also excluded.

Ethical approval

An agreement between the respective heads of the ED and department of Psychiatry 'A' was established. The names of the participants were not mentioned in any form. The cross-institutional review board considered this analysis to be exempt from ethical review. Patients admitted to the ED following a suicide attempt had their medical records prospectively documented. Following clinical stabilisation and before discharge, a psychiatric assessment was performed by psychiatrists of the university hospital Hedi Chaker Sfax. One trained investigator conducted all the interviews.

Procedures and assessment measurements

The 17-item HDRS questionnaire was used in the Tunisian dialect (see supplementary material) and consisted of two parts. The first section included sociodemographic characteristics (i.e. age, gender, education level, marital status, economic income, medical condition and family history of mental health disorder). The second part of the questionnaire was composed of the HDRS scale questions.

HDRS scale

The HDRS rating scale includes 17 items. Eight items are scored by a 5-point Likert scale (0 = absent to 4 = severe). Nine items are scored from 0 to 2. The total score is performed by the sum of the items' scores. According to this scale, the depression severity is classified into four categories: no depression (score 0–9), mild depression (score 10–13), mild-to-moderate depression (score 14–17) and moderate-to-severe depression (score >17). The total HDRS score ranged from 0 to 52 points.

Translation of HDRS

The HDRS was translated from French to the Tunisian dialect through a translation and back-translation process. A bilingual Tunisian translator created the translation from French into the Tunisian dialect. Then, this version was translated back into French by a bilingual translator who was fluent in Arabic and French. Upon completion of this process, an expert panel (including five psychiatrists, one psychologist and one biostatistics specialist) compared the two HDRS translations, focussing on the adequacy of the Tunisian version in terms of the socio-cultural context of the items.

Before testing

Before the testing step, we submitted an experimental version of the scale to a 20-subject sample who were randomly selected and had similar characteristics to the target population (meeting the same eligibility criteria).[16–20]

Sampling

The sample calculation was based on previous studies.[18,21,22] A minimum 'statement/subject' ratio of 1–5 was necessary to validate these new versions; with a sample above 100 subjects.[18,21,22] According to this estimation, to have a 90% power with a 5% margin of error, the current study had to include at least 85 participants.

Statistical analyses

Data reported in the text and tables indicate the mean ± standard deviation for numeric variables and percentages or ranges for categorical variables. To compare dichotomous variables, we used the Pearson Chi-square test. The significance level was a two-sided $P < 0.05$ for all tests. For the validation process, we analysed the psychometric properties of the Tunisian version, including construct validity, internal structural validity and external validity.

Construct validity

For confirming the HDRS construct validity in our sample, a principal component analysis (PCA) analysis was launched for the HDRS items. After the extracted factors were found to be significantly correlated, a varimax rotation was used. The Kaiser–Meyer–Olkin measure of sampling adequacy and Bartlett's test of sphericity were used. Factors with Eigenvalues higher than 0.8 were retained in the model.

Internal validity

This was assessed by Cronbach's alpha coefficient.[18,23,24] To confirm consistency, a coefficient of at least 0.7 was expected (Cronbach & Meehl, 1955).[20,23,25] It is recommended that deletion of any of the items should not increase Cronbach's alpha coefficient. Floor and ceiling effects were reported to assess the distribution of

the responses. The rate of floor and ceiling effects were calculated as the proportion of individuals who obtained the lowest ('never or not at all') and the highest ('very often or very much') scores for any of the items.[26]

External validity

We used the Patient Health Quality-9 (PHQ-9) score to explore the external structure validity of the Tunisian version of the HDRS score. We used Spearman's coefficients.

This external validator (PHQ-9 scale) is a self-administered screening scale. It was developed by Kroenke et al. from the depression module of the Primary Care Evaluation of Mental Disorders (PRIMED).[27] It has been validated in its English and Tunisian versions[28] and consists of 9 items derived from the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-4) classification to assess depression. According to PHQ-9, the diagnosis is as follows: no depression (0–10), mild depression (10–14), moderate depression (15–19) and severe depression (PHQ-9 score >20).

The Receiver-Operating Characteristic (ROC) curve was used to analyse the correlation between the total HDRS score and the presence of depression according to the PHQ-9. The area under the ROC curve was estimated by the method of Hanley and McNeill.[29]

Results

Study population characteristics

This study enrolled 101 participants from 111 patients who were hospitalised for a suicide attempt in the ED. Ten patients were not included because they were aged <18 years. All of the eligible patients were included (i.e. no patients met any exclusion criteria). Table 1 details the sociodemographic and clinical characteristics of the participants. The mean age of the study population was 30.9 ± 11.61 years. Most of the participants were female (69.3%), single (51.5%), had a good socio-economic level (76.2%) and a low level of education (85.1%). Approximately half (53.5%) of the patients who were hospitalised for a suicide attempt lived in an urban area.

The median PHQ-9 score was 11.8 (ranging between 0 and 27). According to this scale, the majority of participants had some form of depression: mild (n = 17; 16.8%), moderate (n = 17; 16.8%) and severe depression (n = 20; 19.8%).

The median HDRS score was 14.9. According to this scale, 44.5% of participants had depression (n = 45). Moderate-to-severe depression was observed in 19.8% of patients (n = 20).

Table 1
Sociodemographic and clinical features of the participants (n = 101).

Characteristic	Value
Sex [n (%)]	
Males	31 (30.7)
Females	70 (69.3)
Age in years [mean ± SD]	30.9 ± 11.6
Marital status [n (%)]	
Single	52 (51.5)
Married	38 (37.6)
Divorced	10 (9.9)
Widowed	1 (1.0)
Economic income [n (%)]	
Good (medium to high)	77 (76.2)
Low	24 (23.8)
Education level [n (%)]	
Illiterate	6 (5.9)
Primary	32 (31.7)
Secondary	54 (53.5)
University	9 (8.9)
Living area [n (%)]	
Urban	54 (53.5)
Rural	47 (46.5)
Medical condition [n (%)]	
Past medical illness	33 (32.7)
Past mental health illness	32 (31.7)
Family history of mental health disorder	18 (17.8)
Suicidal attempt mechanism [n (%)]	
Intoxication	95 (94.0)
Phlebotomy	4 (4.0)
Hanging	1 (1.0)
Drowning	1 (1.0)

Psychometric validity

Internal validity

(1) Internal structure validity: All the items in the questionnaire had a representation index >0.5. These variables were able to explain 58.88% of the variance in this model. A Kaiser-Meyer-Olkin measure of sampling adequacy of 0.839 was found, with a significant Bartlett's test of sphericity ($P < 0.001$). By using a rotated component matrix, these variables could be combined into five factors. The total variance explained by five factors was 64.4% (Table 2). (2) Internal consistency reliability: Cronbach's standardised alpha coefficient was 0.86 for the overall scale. All the items correlated greater than 0.5 (Table 3).

External validity

(1) External structure validity: Significant correlations were found between the total HDRS score and the PHQ-9 score (0.80; $P < 0.0001$) and between the HDRS score and the various items of PHQ-9 ($P < 0.0001$ for each item). These results showed that PHQ-9 items Q2, Q4, Q7, Q8, Q9, and Q10 could predict a high HDRS score.

The ROC curve of the HDRS scale, comparing patients with depression based on HDRS and PHQ9 scales, is shown in Fig. 1. The optimal score was 13.5 according to the ROC curve analysis (Fig. 1). The sensitivity and specificity were good at this cut-off point (80.8% and 91.1%, respectively). The area under the curve was high: 0.93 [0.88–0.97] CI; $P < 0.001$.

Table 2
Varimax rotated matrix of the HDRS.

	Item number	Loading factor
Factor 1		
Loss of weight	16	0.820
Somatic symptoms (gastrointestinal)	12	0.750
Genital symptoms	14	0.697
Somatic symptoms general	13	0.680
Anxiety somatic	11	0.561
Insomnia early	4	0.522
Anxiety-psychic	10	0.499
Factor 2		
Retardation	8	0.763
Insomnia late	6	0.666
Depressed mood	1	0.643
Insomnia middle	5	0.642
Work and interest	7	0.627
Feelings of guilt	2	0.564
Factor 3		
Suicide	3	0.820
Agitation	9	0.528
Factor 4		
Hypochondriasis	15	0.825
Factor 5		
Insight	17	0.852

HDRS, Hamilton Depression Rating Scale.

Table 3
Cronbach's alpha of the different items with the total HDRS score.

Item	Cronbach's alpha coefficient
1	0.891
2	0.866
3	0.772
4	0.813
5	0.794
6	0.912
7	0.898
8	0.903
9	0.718
10	0.961
11	0.964
12	0.901
13	0.909
14	0.911
15	0.862
16	0.939
17	0.794

HDRS, Hamilton Depression Rating Scale.

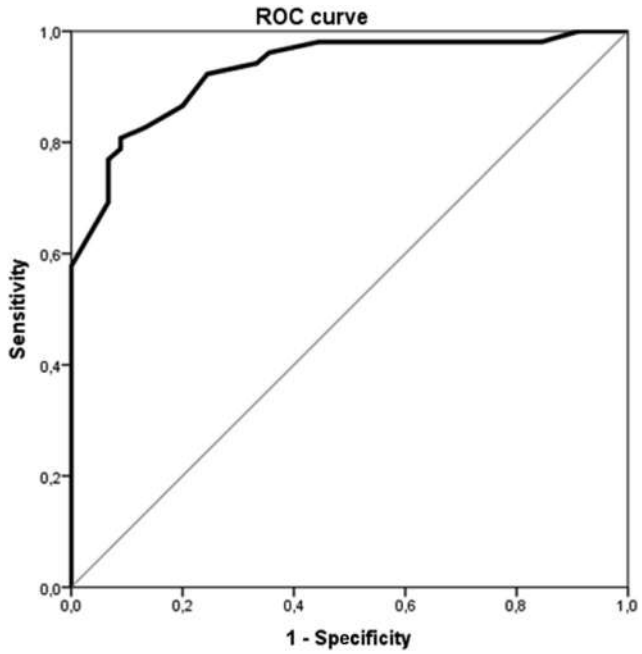


Fig. 1. ROC curve for the relationship between the total HDRS score and the presence of depression, according to PHQ-9. Area under the curve 0.93; Sensitivity: 0.83; Specificity: 0.91; $P < 0.0001$. HDRS, Hamilton Depression Rating Scale; PHQ-9, Patient Health Quality-9; ROC, Receiver-Operating Characteristic.

Discussion

The main objective of this study was to validate the Tunisian dialect version of the HDRS for assessing depression in the context of patients who had attempted suicide. The study results provide initial evidence supporting the reliability and validity of the scale as a screening instrument for Tunisian patients who are hospitalised for a suicide attempt. Depressive disorders, whether isolated or occurring as part of a recurrent mood disorder, were the first diagnosis associated with suicide and suicide attempt.[2,30–32] The HDRS scale, in Turkish, Chinese and Lebanese versions, has been validated in samples of patients with depression.[13–15]

Furthermore, the HDRS scale is significantly correlated with the number of suicide attempts,[33] which explains the choice of our study population (i.e. patients who were hospitalised for a suicide attempt).

Summary of the main descriptive results of the study population

In this study, the majority of participants were female (69.3%), single (51.5%), had a good socio-economic status (76.2%) and a low education level (85.1%). Approximately half (53.5%) of the patients lived in an urban area. The current findings are similar to those of previously published studies.[34,35] There is a current trend for increased suicidal behaviour among young and single individuals, which is related to their psychological immaturity and/or vulnerability[34,36] and lack of social and/or familial support.[35,37]

The demographic and socio-economic characteristics of individuals attempting suicide are different to those who die as a result of suicide.[36,38-41] Women and individuals with low socio-economic status make more suicide attempts, while men and those living in rural areas are more likely to die of suicide.[42] This phenomenon is named the 'gender paradox' in suicide.[23,27,28]

A lack of economic independence could be associated with a higher risk of suicidal behaviour. This is, in part, explained by frustration, leading to interpersonal conflicts and psychological difficulties.[34,43] Similar to other findings, the relationship between a low level of education and suicide attempt was highlighted in the current study.[34,35] A low number of schooling years is regularly associated with low intelligence levels. This may lead to increasing relationship difficulties, which may weaken the individual's balance and defences and increase the risk of suicide.[34]

In the current study, the mean score of PHQ-9 was 11.8. Based on the cut-off scores of PHQ-9, most of the participants had depression: 36.6% had moderate-to-severe depression. Many studies related that suicide attempts are commonly associated with depression.[30,44,45] Ghachem et al.[46] demonstrated that patients with depression have a 30-fold increased risk of suicidal acts compared with individuals without depression.

Psychometric validity

In this study, we were able to validate the Tunisian version of the HDRS. The current results show the reliability and validity of the proposed scale as a screening instrument for patients who have been hospitalised for a suicide attempt in Tunisia. This version of the HDRS scale demonstrated good psychometric properties, with excellent internal consistency for the scale and can, therefore, be used for the assessment of patients who have been hospitalised for a suicide attempt in Tunisia.

Internal validity

In this study, we performed a PCA-type factor analysis with varimax rotation, which allows a suitable representation of the factorial weights. This analysis also expresses the strength or the weakness of correlations of the items.[47] Internal structure validity analyses demonstrated high–grade correlation levels between HDRS individual items and total score.

The total variance illustrated by the five components was 64.4%, which is an acceptable rate.[48] The five components represented the majority of items. Our results are similar to previous studies. The total variance explained by six factors in the Turkish version was 61.3%,[13] in the Lebanese version by four factors was 58.9%,[14] and in the Chinese version by five factors was 52.4%.[15]

Internal consistency reliability: In our study, Cronbach's alpha coefficient was 0.86, which suggests good reliability and internal consistency. Our findings are similar to the Lebanese results and are better than Turkish and Chinese versions[13–15] (Table 4). Thus, we

Table 4
Comparison of the previous studies of HDRS validation.

Study	Year	Country	Population	Cronbach's Coefficient Alpha
Original[11]	1960	UK	Patients with depression	0.84
Chinese[16]	1985	China	Patients with depression	0.71
Turkish[14]	1994–1996	Turkey	Patients with depression	0.75
Lebanese[15]	2017	Lebanon	Patients with depression and controls patients	0.86
The current study	2019	Tunisia	Patients hospitalised for a suicide attempt	0.86

HDRS, Hamilton Depression Rating Scale.

conclude that the Tunisian dialect version of the HDRS has adequate psychometric properties.[49]

External validity

External structure validity: In this current study, the correlation between the HDRS scale and PHQ-9 (which had already been validated in the Tunisian dialect[28]) was tested. The correlation between the total scores was high. The correlation between the HDRS score and the PHQ-9 items was also high. In the Chinese and Turkish studies, the external structural validity of HDRS was demonstrated through significant correlations with the Global Assessment Scale (GAS) [$P < 0.001$][15] and the Clinical Global Impression scale (CGI) [$P < 0.0001$], and the Beck Depression Inventory (BDI) [$P < 0.005$],[13] respectively.

Results of the current study confirm that the proposed Tunisian version of the HDRS is robust and competently assesses depression in patients who have been hospitalised for a suicide attempt.

Finally, the estimated cut-off (13.5) found by ROC curve analysis had good sensitivity (80.8%) and specificity (91.1%). In addition, the properties of the current scale seem to be superior to those of other studies.[14]

Strengths and limitations

The proposed version of the HDRS scale in the Tunisian dialect is simple, understandable, reliable, and shows good external and internal validity. The size of our study population was a strength of the current study because it was comparable to recommendations in the literature.

However, several limitations should be highlighted when interpreting these results. First, this study included data from only one centre; thus, there is the possibility of selection bias. There is a unique dialect in Tunisia, but it would be interesting to test this version of the HDRS in different regions. Second, the study population was among patients hospitalised for a suicide attempt. Future validations in different groups of patients should be conducted before generalising these results. Last, the inter-rater reliability analysis was not performed in the present study.

Despite these limitations, this study will be of interest to physicians and is a base for future studies exploring the proposed Tunisian version of HDRS in the screening of depression. This study will improve the process for the assessment of depression, which is becoming a major public health problem.

Conclusions

In summary, the Tunisian version of the HDRS is an acceptable instrument to screen for depression in patients attempting suicide. This version of the HDRS has good psychometric properties, reliability and internal validity. This study also shows high rates of depression in patients who are hospitalised for a suicide attempt. Further studies should be conducted to generalise these results in different populations.

Author statements

Ethical approval

The cross-institutional review board considered this analysis to be exempt from ethical review.

Funding

All the authors declare they have no funding.

Competing interests

None declared.

Author contributions

Study concept and design (JM); acquisition of the data (NC, RO, RS, IF, BB, ML, FG); analysis of the data (CWO, NC, JJ, JD); drafting of the manuscript (NC, CWO, DM); critical revision of the manuscript (CWO, NR, JM); approval of final manuscript (JD, RN, JM).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2021.11.003.

References

1. OMS. *Journée mondiale de la santé*. WHO; 7 Avril 2017. Available from: https://www.who.int/campaigns/world-health-day/2017/fr/.
2. Ahmed H, Hossain M, Aftab A, Soron T, Alam M, Chowdhury MA, et al. Suicide and depression in the World health organization South-East asia region: a systematic review. WHO South-East asia. *J Public Health* 2017;**6**(1):60–6. https://doi.org/10.4103/2224-3151.206167.
3. World Health Organization. *Preventing suicide. A global imperative*. 2014. ISBN 978 92 4 156477 9, https://apps.who.int/iris/bitstream/handle/10665/131056/9789241564779_eng.pdf;jsessionid=58C82B05F679DAF01B37F549C1DD993E?sequence=1.
4. Ben Khelil M, Gharbaoui M, Farhani F, et al. Impact of the Tunisian Revolution on homicide and suicide rates in Tunisia. *Int J Public Health* 2016;**61**(9):995–1002. https://doi.org/10.1007/s00038-016-0834-8.
5. Mann JJ, Apter A, Bertolote J, et al. Suicide prevention Strategies: a systematic review. *JAMA* 2005;**294**(16):2064–74. https://doi.org/10.1001/jama.294.16.2064.
6. Hawton K, van Heeringen K. Suicide. *The Lancet* 2009;**373**(9672):1372–81. https://doi.org/10.1016/S0140-6736(09)60372-X.
7. Myers JK, Weissman MM. Use of a self-report symptom scale to detect depression in a community sample. *Am J Psychiatry* 1980;**137**(9):1081–4. https://doi.org/10.1176/ajp.137.9.1081.
8. Chipimo PJ, Fylkesnes K. Comparative validity of screening instruments for mental distress in Zambia. *Clin Pract Epidemiol Ment Health* 2010;**6**:4–15. https://doi.org/10.2174/1745017901006010004.
9. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;**23**(1):56–62. https://doi.org/10.1136/jnnp.23.1.56.
10. Bagby RM, Ryder AG, Schuller DR, et al. The Hamilton depression rating scale: has the gold standard become a lead weight? *Am J Psychiatry* 2004;**161**(12):2163–77. https://doi.org/10.1176/appi.ajp.161.12.2163.
11. Poirot I. Syndromes dépressifs et échelles d'évaluation (1 ère partie). *Médecine Sommeil* 2005;**2**(5):38–41. https://doi.org/10.1016/S1769-4493(05)70137-0.
12. Guelfi JD. In: Fabre P, editor. *L'évaluation clinique standardisée en psychiatrie. Boulogne*; 1997. ISBN 13: 9782867870828.

13. Akdemir A, Türkçapar MH, Örsel SD, et al. Reliability and validity of the Turkish version of the Hamilton depression rating scale. *Compr Psychiatr* 2001;**42**(2):161–5. https://doi.org/10.1053/comp.2001.19756.
14. Obeid S, Abi Elias Hallit C, Haddad C, et al. Validation of the Hamilton Depression Rating Scale (HDRS) and sociodemographic factors associated with Lebanese depressed patients. *Encephale* 2018;**44**(5):397–402. https://doi.org/10.1016/j.encep.2017.10.010.
15. Zheng Y, Zhao J, Phillips M, et al. Validity and reliability of the Chinese Hamilton depression rating scale. *Br J Psychiatry* 1988;**152**:660–4. https://doi.org/10.1192/bjp.152.5.660.
16. Vallerand RJ, Halliwell WR. Vers une méthodologie de validation transculturelle de questionnaires psychologiques: implications pour la recherche en langue française. *Can Psychol Can* 1989;**30**(4):662–80. https://www.lrcs.uqam.ca/wp-content/uploads/2017/04/Vers-une-m%C3%A9thodologie-de-validation-trans-culturelles-de-questionnaires-psychologiques.pdf.
17. Brislin RW. Cross-cultural research methods. In: Altman I, Rapoport A, Wohlwill JF, editors. *Environment and Culture*. Boston, MA: Springer US; 1980. p. 47–82. Available from: http://link.springer.com/10.1007/978-1-4899-0451-5_3.
18. Tsang S, Royse C, Terkawi A. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi J Anaesth* 2017;**11**(suppl 1):S80–9. https://doi.org/10.4103/sja.SJA_203_17.
19. Fuentes López E, Morales F. Construction and validation of questionnaire to assess recreational noise exposure in university students. *Noise Health* 2014;**16**(72):292–8. https://doi.org/10.4103/1463-1741.140509.
20. Corbière M, Negrini A, Durand M-J, et al. Development of the return-to-work obstacles and self-Efficacy scale (ROSES) and validation with workers Suffering from a common mental disorder or musculoskeletal disorder. *J Occup Rehabil* 2017;**27**(3):329–41. https://doi.org/10.1007/s10926-016-9661-2.
21. Anthoine E, Moret L, Regnault A, et al. Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures. *Health Qual Life Outcome* 2014;**12**:176. https://doi.org/10.1186/s12955-014-0176-2.
22. Osborne JW. *Best practices in exploratory factor analysis*. 2014. ISBN-13: 978-1500594343.
23. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;**16**(3):297–334. https://doi.org/10.1007/BF02310555.
24. Charter RA. A breakdown of reliability coefficients by test type and reliability method, and the clinical implications of low reliability. *J Gen Psychol* 2003;**130**(3):290–304. https://doi.org/10.1080/00221300309601160.
25. Fermanian J. Validation des échelles d'évaluation en médecine physique et de réadaptation : comment apprécier correctement leurs qualités psychométriques. *Ann Readapt Med Phys* 2005;**48**(6):281–7. https://doi.org/10.1016/j.annrmp.2005.04.004.
26. Dondé C, Haesebaert F, Poulet E, Mondino M, Brunelin J. Validation of the French version of the auditory hallucination rating scale in a sample of hallucinating patients with schizophrenia. *Can J Psychiatr* 2020;**65**(4):237–44. https://doi.org/10.1177/0706743719895641.
27. Spitzer RL. Validation and utility of a self-report version of PRIME-MDThe PHQ primary care study. *JAMA* 1999;**282**(18):1737–44. https://doi.org/10.1001/jama.282.18.1737.
28. Belhadj H, Jomli R, Ouali U, Zgueb Y, Nacef F. Validation of the Tunisian version of the patient health questionnaire (PHQ-9). *Eur Psychiatr* 2017;**41**:S523. https://doi.org/10.1016/j.eurpsy.2017.01.695.
29. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;**143**(1):29–36. https://doi.org/10.1148/radiology.143.1.7063747.
30. Cummins N, Scherer S, Krajewski J, et al. A review of depression and suicide risk assessment using speech analysis. *Speech Commun* 2015;**71**:10–49. https://doi.org/10.1016/j.specom.2015.03.004.
31. Park S, Lee Y, Youn T, Kim BS, Park JI, Kim H, et al. Association between level of suicide risk, characteristics of suicide attempts, and mental disorders among suicide attempters. *BMC Publ Health* 2018;**18**:477. https://doi.org/10.1186/s12889-018-5387-8.
32. Ribeiro JD, Huang X, Fox KR, et al. Depression and hopelessness as risk factors for suicide ideation, attempts and death: meta-analysis of longitudinal studies. *Br J Psychiatry* 2018;**212**:279–86. https://doi.org/10.1192/bjp.2018.27.
33. Desseilles M, Perroud N, Guillaume S, et al. Is it valid to measure suicidal ideation by depression rating scales? *J Affect Disord* 2012;**136**(3):398–404. https://doi.org/10.1016/j.jad.2011.11.013.
34. Mutsapata D M, Dube-Mawerewere V. Factors contributing to parasuicide among patients aged 18 - 35 admitted at a national psychiatric hospital, Zimbabwe. *Open J Nurs* 2017;**7**(8):893–904. https://doi.org/10.4236/ojn.2017.78066.
35. Aarab C, Elghazouani F, Aalouane R, et al. Suivi prospectif sur 5 ans des tentatives de suicide en population clinique dans la région de Fès, Maroc. *Pan Afr Med J* 2014;**18**:321. https://doi.org/10.11604/pamj.2014.18.321.3726.
36. El Majzoub I, El Khuri C, Hajjar K, et al. Characteristics of patients presenting post-suicide attempt to an academic medical center emergency department in Lebanon. *Ann Gen Psychiatr* 2018;**17**:21. https://doi.org/10.1186/s12991-018-0191-5.
37. Choi H-Y, Kim S-I, Yun KW, et al. A study on correlation between anxiety Symptoms and suicidal ideation. *Psychiatry Investig* 2011;**8**(4):320–6. https://doi.org/10.4306/pi.2011.8.4.320.
38. Halayem S, Ounalli H, Boudali M, Hajri M, Abbes Z, Bouden A. Evolution of suicide attempts in a Tunisian clinical population between 2005 and 2015: new modalities for young people to commit suicide? *Encephale* 2018;**44**(6):504–11. https://doi.org/10.1016/j.encep.2017.09.006.
39. Schrijvers DL, Bollen J, Sabbe BGC. The gender paradox in suicidal behavior and its impact on the suicidal process. *J Affect Disord* 2012;**138**(1–2):19–26. https://doi.org/10.1016/j.jad.2011.03.050.
40. Batty GD, Kivimäki M, Bell S, et al. Psychosocial characteristics as potential predictors of suicide in adults: an overview of the evidence with new results from prospective cohort studies. *Transl Psychiatry* 2018;**8**(1):22. https://doi.org/10.1038/s41398-017-0072-8.
41. Nock M, Borges G, Ono Y, editors. *Suicide: global perspectives from the WHO World mental health surveys*. Cambridge: Cambridge University Press; 2012. p. 391 (Cambridge medicine).
42. Cohidon C, Rabet G, Caillet E, Imbernon E. Risque suicidaire et activité professionnelle. *Bulletin Epidémiologique Hebdomadaire* 2011;**47–48**:501–4.
43. Mars B, Burrows S, Hjelmeland H, Gunnell D. Suicidal behaviour across the African continent: a review of the literature. *BMC Publ Health* 2014;**14**:606. https://doi.org/10.1186/1471-2458-14-606.
44. Pompili M, Belvederi Murri M, Patti S, Innamorati M, Lester D, Girardi P, et al. The communication of suicidal intentions: a meta-analysis. *Psychol Med* 2016;**46**(11):2239–53. https://doi.org/10.1017/S0033291716000696.
45. Gili M, Castellví P, Vives M, et al. Mental disorders as risk factors for suicidal behavior in young people: a meta-analysis and systematic review of longitudinal studies. *J Affect Disord* 2019;**245**:152–62. https://doi.org/10.1016/j.jad.2018.10.115.
46. Ghachem R, Boussetta A, Benasr A, Oumaya N. Suicide et pathologie mentale à Tunis : étude rétrospective sur 12 ans à l'hôpital Razi. *Inf Psychiatr* 2009;**85**(3):281–95. https://doi.org/10.3917/inpsy.8503.0281.
47. Cattell R. *The scientific use of factor Analysis in behavioral and life sciences*. New York: Plenum Press; 1978. https://doi.org/10.1007/978-1-4684-2262-7. Online ISBN:978-1-4684-2262-7.
48. Tabachnick BG, Fidell LS. *Using multivariate statistics*. 5th ed. Boston: Pearson/Allyn & Bacon; 2007: 980.
49. Pestana MH, Gageiro JN. Data analysis for social sciences. In: *The complementarity of SPSS*. 6th ed. Lisbonne: Silabo; 2003.