Public Health 221 (2023) 175–180

Contents lists available at ScienceDirect

Public Health

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

Original Research

How important is it to avoid indices of deprivation that include health variables in analyses of health inequalities?

G. McCartney [a, *], R. Hoggett [b], D. Walsh [c], D. Lee [d]

[a] College of Social Sciences, University of Glasgow, 40 Bute Gardens, Glasgow, G12 8RT, UK
[b] NHS Education for Scotland, 89 Hydepark St, Glasgow, G3 8BW, UK
[c] Glasgow Centre for Population Health, Olympia Building, Bridgeton Cross, Glasgow, UK
[d] School of Mathematics and Statistics, Mathematics and Statistics Building, Glasgow, G12 8SQT, UK

ARTICLE INFO

Article history:
Received 7 March 2023
Received in revised form
13 June 2023
Accepted 19 June 2023
Available online 18 July 2023

Keywords:
Deprivation
Health inequalities
Endogeneity
Circular logic
Mathematical coupling

ABSTRACT

Objectives: This study aimed to quantify the difference in mortality inequalities using the Scottish Index of Multiple Deprivation (SIMD) and the Income and Employment Index (IEI; a subindex of SIMD, which excludes health) as ranking measures in Scotland.
Study design: This ecological study was a cross-sectional analysis of routine administrative data.
Methods: Data from the 2020 SIMD and the subindex using data from only the Income and Employment domains, the IEI, were obtained. The correlation between data zones, percentage of data zones that changed deprivation tenth and differences in the Slope Index of Inequality (SII) and Relative Index of Inequality (RII) for Standardised Mortality Ratios (SMRs) across tenths were compared when data zones were ranked by SIMD and IEI.
Results: There was a close correlation between data zones ranked by SIMD and IEI ($R^2 = 0.96$). When data zones were ranked by IEI, 18.7% of data zones moved to a lower deprivation tenth, and 20.8% of data zones moved to a higher deprivation tenth, compared with SIMD. However, only a negligible number of data zones moved two or more tenths. The SMRs across deprivation tenths were very similar between the SIMD and IEI, as were the summary health inequality measures of SII (87.3 compared with 85.7) and RII (0.88 and 0.86).
Conclusion: Although there is a logical problem in using deprivation indices that include health outcomes to rank areas to calculate the scale of health inequalities, the impact of using an alternative subindex containing only data from the income and employment domains is minimal. For population-wide analyses of health inequalities in Scotland, the SIMD does not introduce a substantial bias in the health inequalities summary measures despite substantial movement of small areas between ranked population tenths. Although not examined here, this is likely to be relevant to other similar indices across the United Kingdom.
© 2023 The Author(s). Published by Elsevier Ltd on behalf of The Royal Society for Public Health. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Introduction

Administrative data on deprivation are substantially more readily available in the United Kingdom at a small area level than at an individual or household level. As such, the calculation of health inequality summary measures often relies on indices of area deprivation to produce a population ranking. In these cases, the area deprivation measure is not necessarily theorised as a causal influence but instead as a means of exposing the scale of inequality.[1]

However, many indices of deprivation include health outcomes as part of their weighting, and this introduces a circular logic, or endogeneity bias, in the use of the indices to rank areas for the calculation of health inequalities.[2,3] This problem can be avoided if a subindex is used, which excludes the health measures, but such subindices are not always readily available to researchers. As such, many health inequalities analyses using area deprivation indices as the ranking measure involve an assumption that inclusion of the health domain will have no impact on the results.

The Scottish Government developed the Scottish Index of Multiple Deprivation (SIMD) to support policy-making and analyses of inequalities in Scotland. A weighted measure is calculated for data zones, small areas with a median population of 755, and mean of 778, persons. The SIMD (2020) contains data across seven domains:

* Corresponding author.
E-mail address: Gerard.mccartney@glasgow.ac.uk (G. McCartney).

https://doi.org/10.1016/j.puhe.2023.06.028
0033-3506/© 2023 The Author(s). Published by Elsevier Ltd on behalf of The Royal Society for Public Health. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

income, employment, education, health, access to services, crime and housing.[4] The Income and Employment Index (IEI) uses only the data within those two domains. Data zones are then ranked by their scores and frequently stratified into population tenths or fifths for comparison (noting that deciles and quintiles refer to the points of division between tenths and fifths). This is very similar to the indices used elsewhere in the United Kingdom, including the Index of Multiple Deprivation (IMD, England),[5] Welsh Index of Multiple Deprivation (Wales),[6] and the Northern Ireland Multiple Deprivation Measure (Northern Ireland),[7] all of which include a health domain alongside a collection of other domains similar to Scotland.

This article aims to estimate the change in ranking of data zones when the SIMD is compared with the IEI and what impact a change from SIMD to IEI makes on the measurement of inequalities in mortality in Scotland.

Methods

We obtained SIMD and IEI rankings, mortality data and population sizes from the Scottish Government. Full details of how the SIMD and IEI rankings are calculated are provided elsewhere.[8,9] Briefly, data for 30 indicators across seven domains are collated for each data zone area. The counts for the employment, income, crime and housing domains are divided by the working age or total population sizes as appropriate to create a series of domain scores for each area. For the indicators in the health, education and access domains, the indicators are ranked, transformed to a normal distribution and then combined using factor weights to create domain scores. Each domain score for each area is then standardised and transformed to an exponential distribution and then combined using weights (greatest weight is given to employment and income domains, least to the housing and crime domains) to give an SIMD score for each area (Table 1). Finally, these scores are ranked to give the overall SIMD rank. It is this ranking that is frequently used to determine deprivation fifths or tenths.[7] For the IEI, the income and employment domain scores were transformed exponentially before being combined with equal weighting (as is also the case for those domains within the full SIMD) and then ranked.[8]

We compared the data zone rankings and correlation coefficients for the rankings by SIMD and IEI before calculating the percentage of data zones that changed the deprivation tenth when the SIMD ranking was compared with the IEI ranking.

Using the Standardised Mortality Ratios (SMRs), standardised to the Scottish population, that are part of the SIMD as our health outcome of interest, we calculated the mean SMR within each tenth of data zones ranked by the SIMD index and then by the IEI. Following Pamuk, we then calculated the Slope Index of Inequality (SII) and Relative Index of Inequality (RII) using linear regression methods as per the Scottish Public Health Observatory tool (see https://www.scotpho.org.uk/comparative-health/measuring-inequalities/).[10]

Results

Fig. 1 shows the correlation between the deprivation ranking of data zones using the SIMD and IEI indices. There is a clear linear correlation between the indices with an R^2 of 0.9566, indicating that approximately 96% of the variation in the IEI was explained by the SIMD. There is greater variation between the indices for less deprived areas (i.e. heteroscedasticity), reflecting that we were not using transformed raw scores (of employment or income support benefit claims per unit working age population) for comparison (the data for which are not available publicly). The high correlation is likely to reflect, at least in part, the higher weighting given to the income and employment domains in SIMD.

Despite the high correlation between the indices, a substantial percentage of data zones moved deprivation tenths when the SIMD and IEI indices were compared (Fig. 2). Some 18.7% of data zones became more deprived, and 20.8% became less deprived when the IEI tenth was subtracted from the SIMD tenth, but 60.5% did not

Table 1
Summary of the construction of the Scottish Index of Multiple Deprivation (SIMD).[4]

Domain	Contributing data	Overall weighting[a]
Income	• Number of adults receiving income support (IS), income-based employment and support allowance (ESA), or Jobseeker's' allowance (JSA) • Number of adults receiving Guaranteed Pension Credit • Number of children in IS, JSA or ESA households • Number of people claiming Universal Credit (UC) and their dependent children (excluding those in the 'working with no requirements' conditionality group) • Number of adults and children in Tax Credit families on low incomes	28%
Employment	• Working age recipients of JSA • Working age recipients of Incapacity Benefit (IB), ESA, or severe disablement allowance • Working age recipients of UC not in employment	28%
Health	• Standardised Mortality Ratio • Hospital stays related to alcohol misuse • Hospital stays related to drug misuse • Comparative illness factor • Emergency stays in hospital • Proportion of population being prescribed drugs for anxiety, depression or psychosis • Proportion of live singleton births of low birth weight	14%
Education, skills and training	• School pupil attendance • Attainment of school leavers	14%
Geographic access to services	• Private transport • Travel time to GP surgery, retail centre, petrol stations, primary and secondary schools, post office • Digital access • Proportion of premises unable to receive superfast broadband	9%
Crime	• Recorded crime rates for crimes of violence, sexual offences, domestic house breaking, vandalism, drug offences, and common assault	5%
Housing	• Persons in overcrowded housing • Persons in households without central heating	2%

GP, general practitioner.
[a] Each domain is standardised by ranking the scores and then exponentially transformed before weighting.

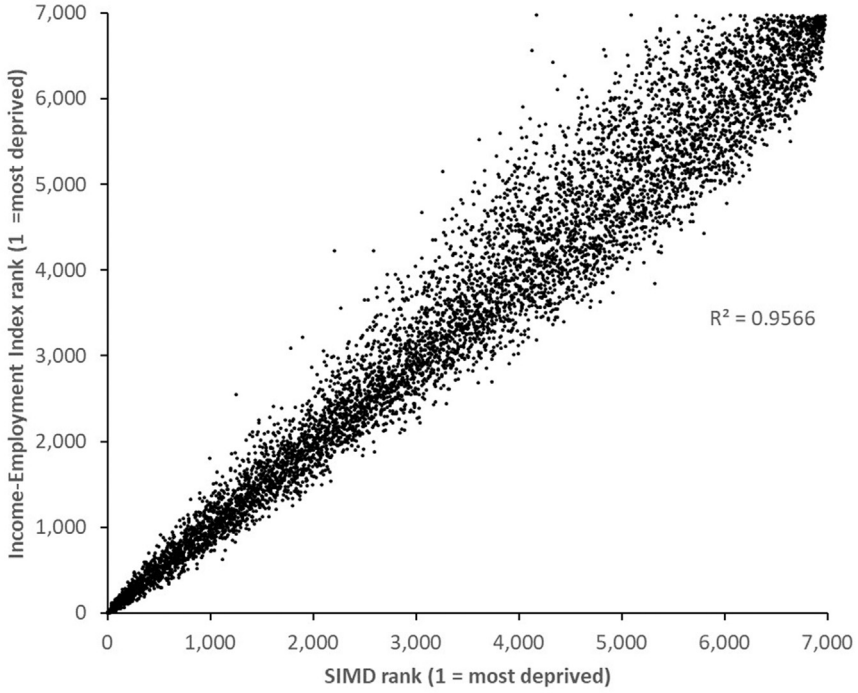

Fig. 1. Correlation between SIMD and IEI deprivation ranks for Scottish data zones, 2020, Scotland.

change, and there was a negligible number that changed two or more tenths. Most of this change was amongst the least deprived data zones (Tables 2 and 3).

The SMRs show reasonably linear gradients across deprivation tenths ranked by both SIMD and IEI, with only small percentage differences in the values (Fig. 3). Again, the variation between the two deprivation indices was greater for the least deprived areas, reflecting the greater movement of data zones between the SIMD and IEI rankings at this end of the range. The SIIs for mortality using SIMD and IEI to rank data zones were 87.3 and 85.7, respectively. The equivalent RIIs were 0.13 for both SIMD and IEI.

Discussion

Main results

Although there is an important theoretical problem with using deprivation indices that include health measures to rank areas for the measurement of health inequalities, this article shows that the use of a narrower index including only the Income and Employment domains in Scotland makes almost no difference to the overall results. This is despite approximately 40% of data zone areas moving deprivation tenth between the full SIMD and the IEI.

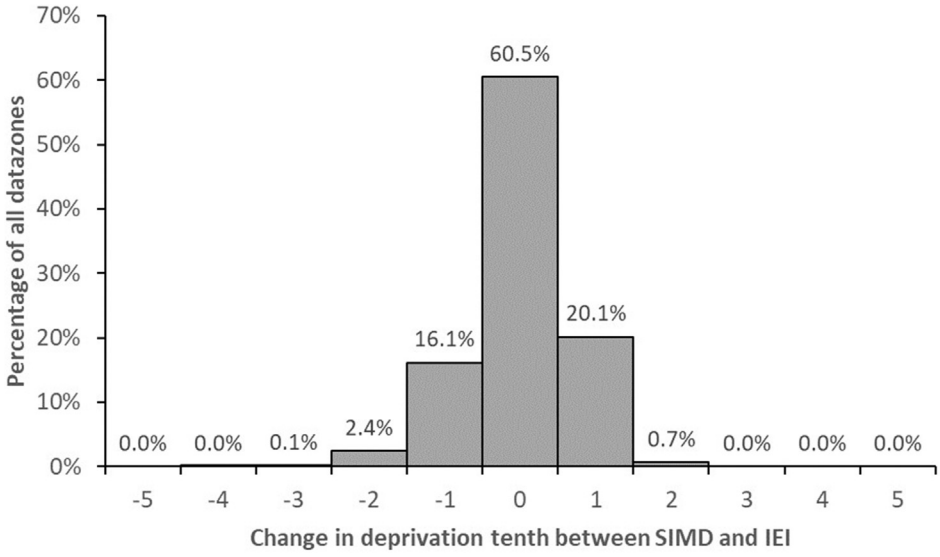

Fig. 2. Change in deprivation tenth for data zones between SIMD and IEI deprivation ranks, 2020, Scotland.

Table 2
Number of data zones in each SIMD and IEI tenth.

IEI tenth (1 is most deprived)	SIMD tenth (1 is most deprived)										Total
	1	2	3	4	5	6	7	8	9	10	
1	646	53	0	0	0	0	0	0	0	0	699
2	52	569	75	0	0	0	0	0	0	0	696
3	0	74	528	96	0	0	0	0	0	0	698
4	0	1	92	468	134	2	0	0	0	0	697
5	0	0	3	120	397	169	9	0	0	0	698
6	0	0	0	11	142	314	217	14	0	0	698
7	0	0	0	2	24	156	286	213	16	0	697
8	0	0	0	0	1	51	144	266	230	6	698
9	0	0	0	0	0	4	38	165	273	217	697
10	0	0	0	0	0	2	3	40	178	473	696
Total	698	697	698	697	698	698	697	698	697	696	6974

IEI, Income and Employment Index; SIMD, Scottish Index of Multiple Deprivation.

Overall, the use of the SIMD for ranking areas for mortality inequality analyses in Scotland seems to make a negligible difference to results, but there is much greater variability for less deprived areas between indices and so analyses that focus on subnational populations or strata should check whether this impacts on results before assuming this is the case.

Strengths and limitations of the analysis

This article uses routine administrative data for both the ranking measure (SIMD and IEI) and the outcome measure (SMR), both of which use data from national data sets without substantive risks of bias. There are other ways of calculating SII and RII using slightly different assumptions about the underlying nature of the data.[11] The findings in this article should thus be checked when other methods are used. Furthermore, the findings may not hold when applied at other spatial levels (e.g. for ranked data zones within local authorities) or when generalised to other data sets (e.g. other nations' deprivation indices or other health outcome datasets), and so caution is urged before applying these findings more broadly. Ideally, we would have replicated these analyses for age-standardised mortality rates, but the numerator and denominator counts in each age group and data zone contain small numbers and are therefore not published routinely. Finally, it is worth noting that even with the IEI domain, this does not entirely remove the circular logic in the deprivation–health relationship, as several of the data sources within those domains relate to the number of people claiming for health- and disability-related benefits.

How this fits with the existing literature

A previous study using the 2004 English IMD investigated the impact of including and excluding the health domain from the ranking measure in the calculation of self-rated health within deprivation fifths and on inequalities across fifths. Similar to our findings, there was a negligible impact both on the prevalence within fifths and the summary regression indices across them.[12] More recently, an analysis of Scottish data, again including and excluding the health domain (from SIMD in this case), found no differences in SII or RII indices for any age or sex group, for mortality, self-rated health, or long-term limiting conditions.[3]

Implications for policy, practice and future research

Many other measures of individual, household and area-based socio-economic position and deprivation are available and capture different and important experiences and mechanisms.[13] It is also important to be clear in any health inequality analyses whether socio-economic data are being used simply for ranking populations and measuring the extent of inequalities or whether the intention is to try to explain differences in health between groups (and/or explain the causes of health inequalities),[1] for example, conflation of the measurement of health inequalities using area-based deprivation indices with place effects and risks ignoring the role of underlying economic relationships between social groups.[14] However, for the routine and rapid measurement of health inequalities in the Scottish context, area-based deprivation measures remain a valid monitoring tool at the population level.[15,16] Area deprivation measures do misclassify a large proportion of the population who are individually income or employment deprived and thereby have smaller inequalities than individually ranked populations.[17] Area-based measures are also subject to the modifiable area-unit problem, whereby the homogeneity/heterogeneity of the area (and the drawing of spatial boundaries), and the size of the areas, makes a substantial difference to the extent of inequalities.

Table 3
Percentage of data zones in each SIMD and IEI tenth.

IEI tenth (1 is most deprived)	SIMD tenth (1 is most deprived)										Total
	1	2	3	4	5	6	7	8	9	10	
1	9.3%	0.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	10.0%
2	0.7%	8.2%	1.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	10.0%
3	0.0%	1.1%	7.6%	1.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	10.0%
4	0.0%	0.0%	1.3%	6.7%	1.9%	0.0%	0.0%	0.0%	0.0%	0.0%	10.0%
5	0.0%	0.0%	0.0%	1.7%	5.7%	2.4%	0.1%	0.0%	0.0%	0.0%	10.0%
6	0.0%	0.0%	0.0%	0.2%	2.0%	4.5%	3.1%	0.2%	0.0%	0.0%	10.0%
7	0.0%	0.0%	0.0%	0.0%	0.3%	2.2%	4.1%	3.1%	0.2%	0.0%	10.0%
8	0.0%	0.0%	0.0%	0.0%	0.0%	0.7%	2.1%	3.8%	3.3%	0.1%	10.0%
9	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.5%	2.4%	3.9%	3.1%	10.0%
10	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.6%	2.6%	6.8%	10.0%
Total	10.0%	10.0%	10.0%	10.0%	10.0%	10.0%	10.0%	10.0%	10.0%	10.0%	100.0%

IEI, Income and Employment Index; SIMD, Scottish Index of Multiple Deprivation.

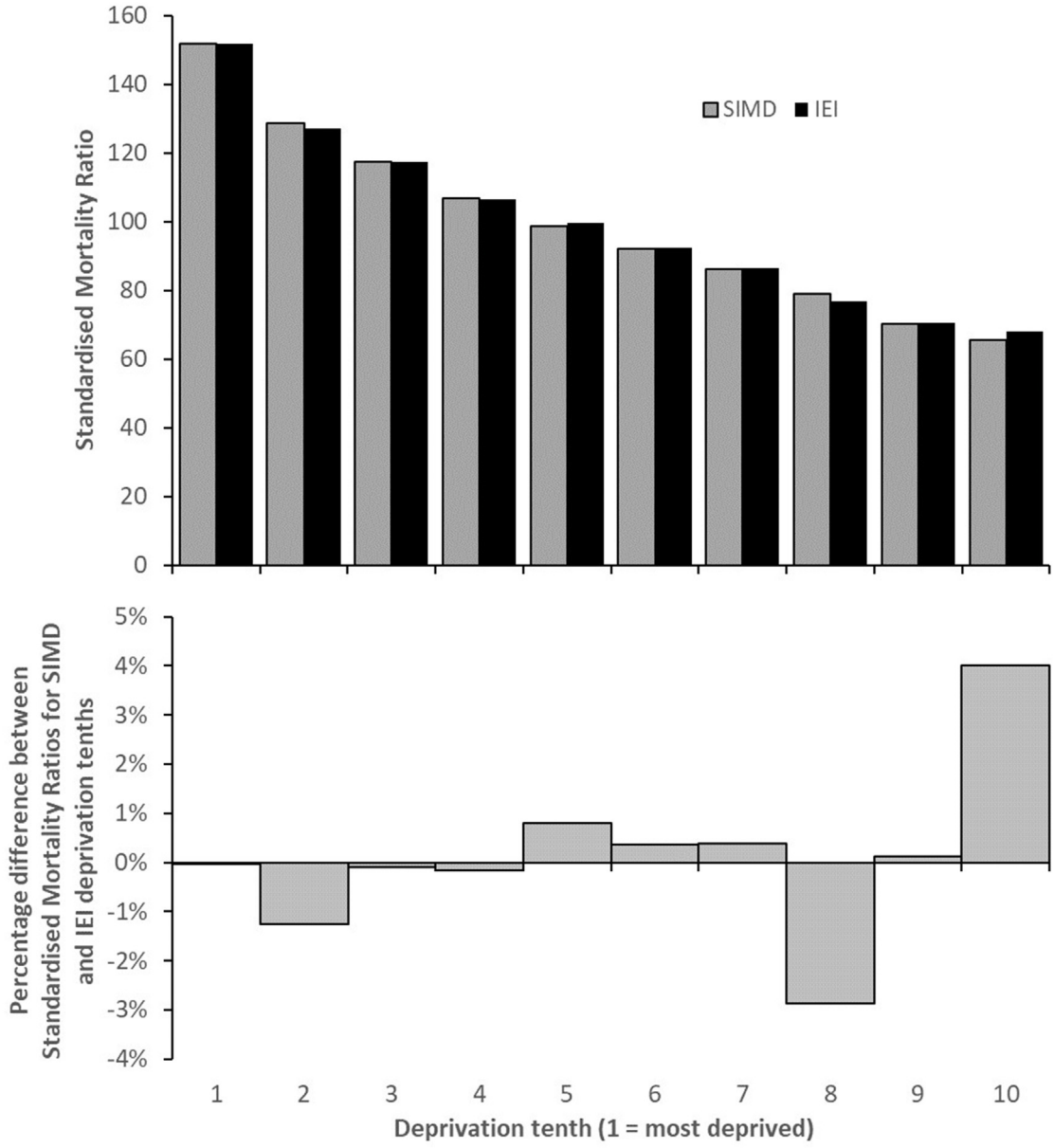

Fig. 3. Difference in Standardised Mortality Ratios between deprivation tenths ranked by SIMD and IEI, 2020, Scotland. IEI, Income and Employment Index; SIMD, Scottish Index of Multiple Deprivation.

However, where the SIMD is used as a ranking measure in the calculation of health inequalities, it makes little or no difference to overall inequality measures of mortality at a national level. It does, however, remain prudent to use area deprivation ranking measures that do not include health outcomes wherever possible. National statistics agencies and governments should ensure that deprivation indices are readily available for each domain separately (and for combined weighted indices such as the IEI) to facilitate such health inequality analyses.

Conclusion

For health inequalities analyses of mortality in Scotland, the use of the full SIMD makes little or no difference to the overall results compared with the use of the IEI. The likelihood of introducing bias in whole population health inequalities analyses using SIMD is small.

Author statements

Ethical approval

No ethical approval was sought or required for this study.

Funding

This study was not funded. G.M., L.E. and D.L. are salaried employees of the University of Glasgow. R.H. and D.W. are salaried employees of NHS Scotland.

Competing interests

All authors declare they have no competing interests.

References

1. McCartney G, Bartley M, Dundas R, Katikireddi SV, Mitchell R, Popham F, et al. Theorising social class and its application to the study of health inequalities. *SSM Popul Health* 2019;**7**:100315. https://doi.org/10.1016/j.ssmph.2018.10.015.
2. McCartney G, Popham F, McMaster R, Cumbers A. Defining health and health inequalities. *Publ Health* 2019;**172**:22–30. https://doi.org/10.1016/j.puhe.2019.03.023.
3. Bradford DRR, Allik M, McMahon AD, Brown D. Assessing the risk of endogeneity bias in health and mortality inequalities research using composite measures of multiple deprivation which include health-related indicators: a case study using the Scottish Index of Multiple Deprivation and population health and mortality data. *MedRxiv* 2022;**12.01**:22282961. https://doi.org/10.1101/2022.12.01.22282961.
4. Scottish Government. *Scottish index of multiple deprivation*. 2020. Available at, https://www.gov.scot/collections/scottish-index-of-multiple-deprivation-2020/. [Accessed 23 August 2021].
5. *English indices of deprivation 2019*. London: UK Government Ministry of Housing, Communities & Local Government; 2019. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019.
6. *Welsh index of multiple deprivation*. Cardiff, Stats Wales; 2019. https://statswales.gov.wales/Catalogue/Community-Safety-and-Social-Inclusion/Welsh-Index-of-Multiple-Deprivation#:~:text=The%20Welsh%20Index%20of%20Multiple,to%201%2C909%20(least%20deprived).
7. *Northern Ireland Multiple deprivation measure*. Belfast: Northern Ireland Statistics and Research Agency; 2017. https://www.nisra.gov.uk/statistics/deprivation/northern-ireland-multiple-deprivation-measure-2017-nimdm2017.
8. *SIMD 2020 technical notes*. Edinburgh: Scottish Government; 2020.
9. *Long-term monitoring of health inequalities: March 2022 report*. Edinburgh: Scottish Government; 2022. https://www.gov.scot/publications/long-term-monitoring-health-inequalities-march-2022-report/documents/.
10. Pamuk ER. Social class inequality in mortality from 1921 to 1972 in England and Wales. *Popul Stud* 1985;**39**(1):17–31.
11. Moreno-Betancur M, Latouche A, Menvielle G, Kunst AE, Rey G. Relative index of inequality and slope index of inequality: a structured regression framework for estimation. *Epidemiology* 2015;**26**(4):518–27.
12. Adams J, White M. Removing the health domain from the Index of Multiple Deprivation 2004—effect on measured inequalities in census measure of health. *J Publ Health* 2006;**28**(4):379–83. https://doi.org/10.1093/pubmed/fdl061.
13. Davey Smith G, Hart C, Watt G, Hole D, Hawthorne V. Individual social class, area-based deprivation, cardiovascular disease risk factors, and mortality: the Renfrew and Paisley study. *J Epidemiol Community Health* 1998;**52**:399–405.
14. Sayer A, McCartney G. Economic relationships and health inequalities: improving public health recommendations. *Publ Health* 2021:103–6.
15. Frank J, Haw S. Best practice guidelines for monitoring socioeconomic inequalities in health status: lessons from Scotland. *Milbank Q* 2011;**89**(4):658–93.
16. McCartney G, Leyland AH, Fischbacher C, Whyte B, Walsh D, Stockton D. Long-term monitoring of health inequalities in Scotland: a response to Frank and Haw. *Milbank Q* 2013;**91**(1):187–92.
17. McCartney G, Hoggett R, Walsh D, Lee D. How well do area-based deprivation indices identify income- and employment-deprived individuals across Great Britain today? *Publ Health* 2023;**217**:22–5. https://doi.org/10.1016/j.puhe.2023.01.020.