PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Changing mortality trends in countries and cities of the United Kingdom (UK): a population-based trend analysis.
AUTHORS	Walsh, David; McCartney, Gerry; Minton, Jon; Parkinson, Jane; Shipton, Deborah; Whyte, Bruce

VERSION 1 – REVIEW

REVIEWER	Rosie Seaman University of Stirling, Scotland
REVIEW RETURNED	06-Mar-2020

GENERAL COMMENTS	This is a clearly written and timely paper. The authors have done an excellent job in co-ordinating so many data sources and presenting this volume of results so clearly and coherently. I have two minor points that I would like a little more clarification on before publication. 1. The fact that some of the ICD cause of death categories are not mutually exclusive seems potentially problematic. Does this mean the same death could be counted twice in the results? This could be my misunderstanding. 2. The Scottish city specific results - what is the value of presenting the IMD results as city specific deprivation? Please provide some information on what the authors actually did to the data, why this is important, and why it was only done for Scotland. I would argue that the measures of multiple deprivation are intended to be used to reflect deprivation in a consistent way at the national level for each country and readjusting them to be city specific seems to go against the intended use.

REVIEWER	Lawrence Best University College London UK
REVIEW RETURNED	19-Apr-2020

GENERAL COMMENTS	Thank you for this important piece of work. It is an issue which is critical to healthcare provisioning. In general the paper is well done, I picked up on a couple typographical errors which I've listed at the bottom, there may be others I missed. In terms of substantive points, I don't think it is overly clear how the measures of inequality are derived or what they actually mean. It would be worth clarifying this, both in the abstract and in the main text. Aside from this I think it would be worth calculating the excess deaths caused by the reduction in improvement of mortality rates. You could simply do this by averaging the improvement from previous time-periods and

	extrapolating. I know this is somewhat simplistic and will just give a number which may or may not be meaningful but it will again simplify the message of the paper. This is more a suggestion than anything else so feel free to ignore if you don't think it is appropriate. If you really wanted to reinforce the message you could correlate the worsening improvement with funding for various health services, get a r value etc. as currently linking it to austerity is slightly spurious (though I do agree it is the likely cause). Aside from these points I believe the paper adds value to the current literature and is methodologically and literarily rigorous enough for publication. Typographical errors: Bottom of page 7 "The aim was to compared". Page 14 some reference numbers not superscripted

REVIEWER	Silvia Rizzi Interdisciplinary Center on Population Dynamics Department of Epidemiology Biostatistics and Biodemography University of Southern Denmark Denmark
REVIEW RETURNED	27-May-2020

GENERAL COMMENTS	The article is easy to read, generally well written and structured. It is an interesting study that adds to the literature a detailed overview of mortality trends over ca. 40 years in subparts of the UK (England and Wales, Scotland and Northern Ireland, and main cities). Major comments: a. The worsening of mortality trends both for males and females, and particularly for the most deprived group, due to drug-related poisoning is a key finding. I suggest to calculate those gender and deprivation specific rates by age class as well. b. Make sure that title, abstract and results highlight the key findings. The substantial worsening of mortality trends occurs for the most deprivated; for external causes and drug-related poisonings; mostly in Scotland. c. There is a discrepancy between what described in the Method Section of the Abstract and the Methods/Results description in the main text: The indicators (RII and SII) mentioned in the abstract are not addressed enough in the text. I suggest extending the Statistical analyses with a clear explanation of the computed indicators and the Results with an interpretation of the corresponding outcomes. Detailed comments: 1. In the Abstract do not use abbreviations: the indicators RII and SII of the Methods should be written extensively (page 3, line 33). 2. In the Geography part (page 8, line 9-29) mention briefly the variables used to construct the indices of deprivation of the registries. 3. In the Statistical analyses (page 8, line 39) make the age-groups used clearer and justify the chosen age classes. Additionally, why infant mortality (0-1 year) and/or mortality at younger ages (1-14 years) is not reported? 4. In the Statistical analyses (page 8, line 46) “between 1981/83 and 1985/88”, the latter 88 should be replaced by 87. 5. In the Statistical analyses (page 9, line 15-17) give an overview of

	the indexes used. How were they calculated? Which assumptions have to be met? How can one interpret the results? Please extend to make this clear to the reader. 6. At the end of the Statistical analyses state which software was used for the statistical analysis.

REVIEWER	Leigh Blizzard Menzies Institute for Medical Research Australia
REVIEW RETURNED	17-Jul-2020

GENERAL COMMENTS	I have some concerns: 1. The data are disguised as moving averages. I can understand why this has been done. By smoothing fluctuations and making trends easier to discern, this no doubt greatly improves the visual presentation of the data in Figures 1, 3 and 4 (and Web Figures 2, 4–7). The results are beautiful, but a tarradiddle from the viewpoint of the scientific method. Figure 2 contains the comparisons that could be important, but it too is based on moving averages compares overlapping periods of time. For example, the collection 2011/13–2015/17 overlaps the overlapping collection 2006/08–2010/12. In consequence, no statistical testing is possible and the comparison is meaningless; 2. What rationale is there for selecting 2012–16 as the key epoch in Figure 2? (The epoch 2012–16 corresponds to the collection 2011/13–2015/17). What are the events and circumstances of 2012–16 that make it so remarkable? Without a strong case being made for the occurrence of differentiating events and circumstances during 2012-16, its selection is suspicious. The authors could have chosen 2013–17. Better still, they could have chosen 2011–15 and analysed the time series in the standard sequence 1981–85, 1986–90, 1991–95 …, 2006–2010, 2011–15 and thereby avoided all suspicion of cherry-picking; 3. The findings do not justify the conclusion that "It is imperative that a range of policies are introduced at UK Government level to reverse previous cuts to social security and social services, and to therefore protect the health of the most vulnerable in society." The authors have not demonstrated that their "results" stand up to statistical testing. Without that, they play into the hands of the "alternative facts" cabal who claim that scientists deliberately inflate the dangers of their favourite issues. Were the manuscript to be published in its present form, I am concerned that the dark forces of the fact-free world would leap on it as bad science. But a revised manuscript could make a useful contribution. One option would be to make fundamental changes to it to adopt a more scientific approach. Use standard 5-year periods, apply tests of trend with frequentist or Bayesian change-point analysis, etc. Another option would be to change the focus of the manuscript to make it a harbinger of future events, by providing an "early warning" that the rate of decline in mortality rates may have diminished. Something like: Warning – there are reports around the world that the rate of decline in mortality has diminished – this may have happened in the UK – recent data for the UK show some signs of this – it is important to continue to monitor the developing situations – watch this space.

	More attention needs to be given to phraseology. As a very small sample: • tables that are headed “European age-standardised mortality rates”, and vertical axes that are labelled “ASRM” – the rates reported are for the UK, not for Europe; • frequent misuse of the word “trend” as in “Figure 1 presents trends in all-cause standardised mortality rates” – no, Figure 1 depicts all-cause standardised mortality rates (and strictly it depicts moving averages of all-cause standardised mortality rates); • “flattening, or worsening”; • “five-year periods” that cover 7 years (for example, the seven-year collection 2011/13–2015/17); In summary, the evidence that a change in trend in mortality rates has occurred is not overwhelming. It remains possible that the phenomenon that the authors have identified – arguable by cherry-picking time periods and disguising the real data with moving averages – may resolve itself in future years.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1	
Reviewer Name: Rosie Seaman	
This is a clearly written and timely paper. The authors have done an excellent job in co-ordinating so many data sources and presenting this volume of results so clearly and coherently. I have two minor points that I would like a little more clarification on before publication.	We are grateful to the reviewer – and indeed all the reviewers – for the helpful comments.
1. The fact that some of the ICD cause of death categories are not mutually exclusive seems potentially problematic. Does this mean the same death could be counted twice in the results? This could be my misunderstanding.	As stated in both the Methods section, and as an explanatory note in the relevant online appendix, there are indeed some overlaps in causes of death e.g. all cancers and lung cancer, and the very broad category of external causes and more specific causes such as suicide and MVTAs. We would argue that the inclusion of such a broad set of different causes of death is beneficial rather than problematic, even with such overlaps. For example, it is useful to examine all malignant cancers as one relevant group of deaths, but it is also instructive to focus on more specific cancers such as lung cancer, given the latter’s particular aetiology. Similarly, all external causes represent a broad set of ICD codes that is of interest in its own right as it distinguishes that type of mortality from – for example – deaths from chronic diseases. However, it is

	also obviously of interest to focus on particular sub-sets of external causes such as suicide. Similar analyses of overlapping causes have been the subject of a great many other research papers. Thus we would argue that this is a useful approach, although we agree that it needs to be clearly explained. To that end, we have added further explanatory text below Figure 4 noting this (it reads: 'Note that there are some overlaps in the ICD code definitions for some causes of death e.g.: all cancers and lung cancer; all external causes, intentional self-harm and drug-related poisonings; external causes and MVTAs'). We trust this is all acceptable to the reviewer.
2. The Scottish city specific results - what is the value of presenting the IMD results as city specific deprivation? Please provide some information on what the authors actually did to the data, why this is important, and why it was only done for Scotland. I would argue that the measures of multiple deprivation are intended to be used to reflect deprivation in a consistent way at the national level for each country and readjusting them to be city specific seems to go against the intended use.	Thanks to the reviewer for raising this. The value of this type of analysis is in being able to present a more accurate picture of *within-city* inequalities. This is particularly valuable for a place like Glasgow, where approximately *half the total population* of the city are classed as being in the most deprived *national quintile*[1]. Thus, using nationally-derived quintiles for that purpose is quite unhelpful. Indeed, given the benefits of this approach in terms of a more insightful analysis of inequalities within cities, these city-specific quintiles (and deciles) are published routinely by ISD Scotland (now part of Public Health Scotland) here: https://www.isdscotland.org/Products-and-Services/GPD-Support/Deprivation/SIMD/. This was the source of the data used for these particular analyses. Unfortunately similar data were not available for the English cities. We have added text to the Methods section to clarify both the source of the Scottish data, and the unavailability of them for English urban areas. Again, we hope this satisfies the reviewer's concern.
Reviewer: 2	
Reviewer Name: Lawrence Best	
Thank you for this important piece of work. It is an issue which is critical to healthcare provisioning.	Thanks to the reviewer for his extremely helpful comments.

[1] See https://www.understandingglasgow.com/indicators/poverty/deprivation (figures relate to 2016 SIMD index, one of the indices used in the analyses)

In general the paper is well done, I picked up on a couple typographical errors which I've listed at the bottom, there may be others I missed. In terms of substantive points, I don't think it is overly clear how the measures of inequality are derived or what they actually mean. It would be worth clarifying this, both in the abstract and in the main text. Aside from this I think it would be worth calculating the excess deaths caused by the reduction in improvement of mortality rates. You could simply do this by averaging the improvement from previous time-periods and extrapolating. I know this is somewhat simplistic and will just give a number which may or may not be meaningful but it will again simplify the message of the paper. This is more a suggestion than anything else so feel free to ignore if you don't think it is appropriate. If you really wanted to reinforce the message you could correlate the worsening improvement with funding for various health services, get a r value etc. as currently linking it to austerity is slightly spurious (though I do agree it is the likely cause).	Many thanks for this. In response, we have added text to both the abstract and methods to better explain the calculation and meaning of the measures of inequality (RII and SII). Abstract word count limits mean we have only been able to define (and spell in full) the measures, but we have supplied more detail to the methods section of the main part of the paper, and have additionally highlighted an illustrative example in the results section. The suggestions for potential further analyses are helpful. However, given the volume of analyses already presented in the paper, we were not sure yet more analyses in an appendix would be particularly beneficial. Furthermore, similar research based on comparisons with projections has already been published by two of the co-authors (albeit only at country level and based on life expectancy rather than excess numbers of deaths) - see Minton et al 2020[2]; some of the other suggested analyses such as associating changes to mortality with cuts to services etc have also been undertaken by others, and are cited in our paper. However, in line with the comment regarding spuriousness, and similar points made by the editor and another reviewer, we have amended the text in both the abstract and the main paper regarding the role austerity. We hope these are satisfactory responses to your highly pertinent points.
Aside from these points I believe the paper adds value to the current literature and is methodologically and literarily rigorous enough for publication.	Many thanks.
Typographical errors:	
Bottom of page 7 "The aim was to compared".	Many thanks for spotting this. It has now been corrected.
Page 14 some reference numbers not	Thanks for noting this. I think it's actually some sort of

[2] Minton J, Fletcher E, Ramsay J, Little K, McCartney G. How bad are life expectancy trends across the UK, and what would it take to get back to previous trends?. J Epidemiol Community Health. 2020;74(9):741-746.

superscripted	system error – i.e. this is not a problem in the version of the Manuscript I have in Word. Hopefully any remaining problems will be picked up by editing staff prior to (hopefully!) publication.
Reviewer: 3	
Reviewer Name: Silvia Rizzi	
The article is easy to read, generally well written and structured. It is an interesting study that adds to the literature a detailed overview of mortality trends over ca. 40 years in subparts of the UK (England and Wales, Scotland and Northern Ireland, and main cities).	As with the other reviewers, we are genuinely very grateful for your time and helpful comments.
Major comments:	
a. The worsening of mortality trends both for males and females, and particularly for the most deprived group, due to drug-related poisoning is a key finding. I suggest to calculate those gender and deprivation specific rates by age class as well.	Many thanks for this comment. We have added two additional figures to the online appendix (Web Figures 8 and 9) in an attempt to address this. Web Figure 8 shows drug related poisonings in Scotland, England & Wales and Northern Ireland, by sex, for all ages, 15-44 years and 45-64 years; Web Figure 9 shows the data for 0-64 years by sex and the least and most deprived quintiles for Scotland. Given the age profile of deaths at this age, similar deprivation trends will be apparent for ages 15-44 years and 45-64 years. We have added detail of what these trends show in the main text. We hope this sufficiently addresses the reviewer's point.
b. Make sure that title, abstract and results highlight the key findings. The substantial worsening of mortality trends occurs for the most deprivated; for external causes and drug-related poisonings; mostly in Scotland.	As highlighted above in the response to the editorial comment, we have added further detail to the results section of the abstract. Within the severe constraints of the abstract word count limit, the abstract results section now covers: the changing mortality trends at city and country level (mentioning all age groups); the increased mortality among the most deprived populations (in all countries and

	cities analysed, and for most causes of death); the associated widening of absolute and relative inequalities; the particular issues in Scotland e.g. higher and increasing drug-related deaths. As already stated, additional text (and web figures) have been added to the included results. We hope this all satisfies your concerns.
c. There is a discrepancy between what described in the Method Section of the Abstract and the Methods/Results description in the main text: The indicators (RII and SII) mentioned in the abstract are not addressed enough in the text. I suggest extending the Statistical analyses with a clear explanation of the computed indicators and the Results with an interpretation of the corresponding outcomes.	Thanks for this. We accept that this should have been covered in more detail but were limited in our ability to do so because of the abstract word count limit. We have now added additional text to both the abstract and main text (in both the methods and results sections) which, we hope, satisfies your concerns.
Detailed comments:	
1. In the Abstract do not use abbreviations: the indicators RII and SII of the Methods should be written extensively (page 3, line 33).	This has been changed, as requested.
2. In the Geography part (page 8, line 9-29) mention briefly the variables used to construct the indices of deprivation of the registries.	We have added more detail of the composition of the different deprivation measures used which we hope is satisfactory.
3. In the Statistical analyses (page 8, line 39) make the age-groups used clearer and justify the chosen age classes. Additionally, why infant mortality (0-1 year) and/or mortality at younger ages (1-14 years) is not reported?	Again, thanks for these helpful observations. We have added text to justify the selection of the age groups, and now also show trends for 0-14 years in Web Figure 4. We also refer to this added figure in the results section, and also at the end of the Discussion. We accept that the lack of data on infant mortality is a potential weakness and you are correct to highlight this. Our principal justification for this is that a major focus of the project was city-based mortality and we knew numbers of infant deaths would be very small in such areas. Thus we did not collect infant mortality data for N. Ireland or England & Wales as part of the project. In hindsight, this is an omission and we have now highlighted it as such in the 'strengths & weaknesses' section of the Discussion, where

	we have also referred to the relevance of other work which has examined recent infant mortality trends in England. We hope the reviewer can accept this compromise.
4. In the Statistical analyses (page 8, line 46) “between 1981/83 and 1985/88”, the latter 88 should be replaced by 87.	Many thanks for spotting this typo. It has now been corrected.
5. In the Statistical analyses (page 9, line 15-17) give an overview of the indexes used. How were they calculated? Which assumptions have to be met? How can one interpret the results? Please extend to make this clear to the reader.	As stated, further details have now been added to this section and we hope this satisfies the reviewer’s concern.
6. At the end of the Statistical analyses state which software was used for the statistical analysis.	The detail has been added.
Reviewer: 4	
Reviewer Name: Leigh Blizzard	

1. The data are disguised as moving averages. I can understand why this has been done. By smoothing fluctuations and making trends easier to discern, this no doubt greatly improves the visual presentation of the data in Figures 1, 3 and 4 (and Web Figures 2, 4–7). The results are beautiful, but a tarradiddle from the viewpoint of the scientific method. Figure 2 contains the comparisons that could be important, but it too is based on moving averages compares overlapping periods of time. For example, the collection 2011/13–2015/17 overlaps the overlapping collection 2006/08–2010/12. In consequence, no statistical testing is possible and the comparison is meaningless; 2. What rationale is there for selecting 2012–16 as the key epoch in Figure 2? (The epoch 2012–16 corresponds to the collection 2011/13–2015/17). What are the events and circumstances of 2012–16 that make it so remarkable? Without a strong case being made for the occurrence of differentiating events and circumstances during 2012-16, its selection is suspicious. The authors could have chosen 2013–17. Better still, they could have chosen 2011–15 and analysed the time series in the standard sequence 1981–85, 1986–90, 1991–95 …, 2006–2010, 2011–15 and thereby avoided all suspicion of cherry-picking;	We have responded to comments 1 and 2 together below (as they are obviously connected). The principal reason for smoothing the data by means of rolling averages was not to 'disguise' them, but to make changes over time more easily understood and interpreted by the reader. This is particularly important because of the fluctuation in rates at city level, an issue we highlight in the manuscript. Such fluctuations in the data due to relatively small numbers (particularly when examining particular age groups and/or particular causes of death) can be problematic and an obstacle to interpretation. It is also important to have a consistent approach to reporting the results of the analyses, which is part of the reason for extending the approach of rolling averages to the other, national level, analyses. However, the principal motive is, as stated, for ease of interpretation. It is also important to emphasise that this is an approach that has been taken in countless other pieces of published, valuable (not 'meaningless'), research. In terms of the percentage change in rates between time periods, we also show this for cities (Web Figures 1 and 3) and would argue therefore that given the fluctuation in rates at city level, it is actually quite important <u>not</u> to compare between single years i.e. in case one selected year is an aberration. Examining the difference between three-year periods is much more robust in that sense. And this is why this method was chosen and was extended to the analyses shown in Figure 2. We have inserted additional text to the Methods section to better explain this ("Three-year averages were used to overcome the issue of fluctuating rates (especially at city level)."). Nonetheless, in the hope of assuaging the reviewer's concerns, and certainly in the hope of rebutting the accusation of "cherry-picking" (an accusation we feel is entirely unjust), I have redone the country level analyses that are shown in Figure 2 (the particular focus of his criticism) using the 'standard' five year periods he has suggested i.e. 1981–85, 1986–90, 1991–95 …, 2006–2010, 2011–15. These are shown below, and the results are clearly very similar to those presented in Figure 2 in the manuscript: in fact, if anything, they highlight the change in rates in the last period more clearly, especially for females. However, given that (a) this approach omits the last two years of data (limiting the time period to 2015 rather than 2017) and (b) the importance of using three-

	year periods for cities (as explained above), we would argue that the original Figure 2 should remain in the paper. We have, however, added a comment to the relevant part of the results section to highlight the fact that analysing the data in this alternative manner produced similar results. In relation to the reviewer's second point, the reason for looking at the three year period around 2012 is that, as stated in the introduction, and as highlighted in the results section in relation to Figure 1, mortality trends in the UK appear to have changed **from 2012 onwards**. It's important to emphasise that this is not based solely on our own analyses shown in Figure 1 (about which the reviewer is obviously unsure), but those of other authors who are cited in the paper e.g. Fenton et al (2019) who tested the timing of mortality changes using one-break and two-break segmented regression models. Therefore it obviously makes sense to examine that period in comparison with previous periods to see if that change is also observed for all the geographically-defined areas that are the focus of the paper (i.e. the different countries and cities). The selection of time periods was built around that, and we would argue it is justifiable on the basis of both the existing evidence, and the trends presented in our paper.

3. The findings do not justify the conclusion that "It is imperative that a range of policies are introduced at UK Government level to reverse previous cuts to social security and social services, and to therefore protect the health of the most vulnerable in society." The authors have not demonstrated that their "results" stand up to statistical testing. Without that, they play into the hands of the "alternative facts" cabal who claim that scientists deliberately inflate the dangers of their favourite issues. Were the manuscript to be published in its present form, I am concerned that the dark forces of the fact-free world would leap on it as bad science. But a revised manuscript could make a useful contribution. One option would be to make fundamental changes to it to adopt a more scientific approach. Use standard 5-year periods, apply tests of trend with frequentist or Bayesian change-point analysis, etc. Another option would be to change the focus of the manuscript to make it a harbinger of future events, by providing an "early warning" that the rate of decline in mortality rates may have diminished. Something like: Warning – there are reports around the world that the rate of decline in mortality has diminished – this may have happened in the UK – recent data for the UK show some signs of this – it is important to continue to monitor the developing situations – watch this space.	As described above in response to the editorial comment, the conclusion in the abstract has been changed. We have also amended the conclusion in the main part of the manuscript (to a more general point about introducing policies 'to protect the health of the most vulnerable in society') in the hope that this satisfies the reviewer's point here. More generally, we hope that by means of both the arguments and justifications outlined above, along with the results of the reviewer-suggested analyses of alternative five year periods, that we have addressed the reviewer's principal concerns?

More attention needs to be given to phraseology. As a very small sample:	
• tables that are headed "European age-standardised mortality rates", and vertical axes that are labelled "ASRM" – the rates reported are for the UK, not for Europe;	We are slightly confused by this point. The vertical axes in question are labelled 'EASR' as an abbreviation for 'European age-standardised rate' – because we used the WHO European Standard population for the standardisation. This is best practice, and this population has been used in countless analyses of UK data (indeed they are used by the main Scottish, English and N. Irish national statistical agencies in their official publications). No axes are labelled 'ASRM' and there are no tables with such a heading. We apologise if the reviewer has spotted something that we have missed!
• frequent misuse of the word "trend" as in "Figure 1 presents trends in all-cause standardised mortality rates" – no, Figure 1 depicts all-cause standardised mortality rates (and strictly it depicts moving averages of all-cause standardised mortality rates); • "flattening, or worsening"; • "five-year periods" that cover 7 years (for example, the seven-year collection 2011/13–2015/17);	We have tried to amend a number of sections of the manuscript where the reviewer thinks inappropriate language has been used (e.g. the use of the word 'trend'). However, we would argue that descriptions such as "flattening" and "worsening" make sense to most readers, and we have added text to the Methods section to explain the use of 'five-year' ("For simplicity, we use the expression 'five-year' interval to reflect the mid-points of the three-year average (e.g. 1982 to 1986 in relation to 1981/83 to 1985/87).").
In summary, the evidence that a change in trend in mortality rates has occurred is not overwhelming. It remains possible that the phenomenon that the authors have identified – arguable by cherry-picking time periods and disguising the real data with moving averages – may resolve itself in future years.	We are grateful to the reviewer for his time in reviewing the paper.

VERSION 2 – REVIEW

REVIEWER	Lawrence Best Division of Surgery and Interventional Science University College London
REVIEW RETURNED	23-Sep-2020

GENERAL COMMENTS	All of my comments have been satisfactorily addressed. The comments of other reviewers also appear to have been addressed thoroughly, however that is for the editor and other reviewers to determine.

	From my perspective this manuscript is now ready for publication.

REVIEWER	Silvia Rizzi Interdisciplinary Center on Population Dynamics, University of Southern Denmark, Denmark
REVIEW RETURNED	16-Sep-2020

GENERAL COMMENTS	Comments satisfactorily addressed.