PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	How much evidence is there that political factors are related to population health outcomes? An internationally comparative systematic review
AUTHORS	Barnish, Max; Tørnes, Michelle; Horne, Becky

VERSION 1 – REVIEW

REVIEWER	David RM Smith Public Health England, UK
REVIEW RETURNED	16-Dec-2017

GENERAL COMMENTS	The authors present a worthwhile update to an existing meta-analysis of the influence of political factors on population health outcomes. Although its conclusions are mostly similar, this update has the benefits of more than doubling the number of included studies and assessing the risk of bias across all included studies. However, there are important problems with the way the study is presented. Major: The paper would benefit from a clear, neutral introduction to the relationship between political systems and healthcare outcomes, as opposed to the current (rather jumbled) introduction to historical and ideological links. Further, it is obvious that the authors believe that left-wing policies improve health, and are worried about the health consequences of burgeoning right-wing policies and events. They later show that there is evidence-based precedence for these beliefs, but this is a fundamental outcome of the meta-analysis. Introducing this as a foregone conclusion in the introduction undermines the authors' objectivity and questions the need for this study in the first place. Second, the importance of economic development and political history on the generalizability of these results is unclear. To what extent did included studies control for relative social wealth/development? For example, Cereseto & Waitzkin (1986) conclude that leftist policies have positive effects on health within strata of economic development, but that rightist policies may be linked to greater economic development. This is of particular concern with respect to the variable 'political tradition', in which rich/OECD countries are overwhelmingly (but unsurprisingly) over-represented. But what about elsewhere? The capitalisation of China has entailed enormous economic growth – have healthcare outcomes there not improved substantially as a result? Conversely, the pink tide has recently wrought ruin on several once thriving South American economies – have healthcare systems there not

	deteriorated? Of course the authors are only working with the evidence they could find, but this kind of global context is important to discuss, especially since the finding that left-wing political traditions always entail better health outcomes is largely based on the recent history of social democracies in the wealthy West. Minor: The authors emphasise an internationally comparative approach, but (i) less developed and non-Western nations are underrepresented, and (ii) the narrative focus is clearly on the UK. Perhaps this can be acknowledged, and more global narrative examples could be used, and not just those relevant to Scotland/UK? The tables assessing risk bias are unwieldy. Offering the colour key in the caption and repeating column headings on each page would help the reader. Table 3 serial number 3: Data are from 1960-1994, not 1960-1964

REVIEWER	Natalia Calanzani University of Edinburgh, Scotland
REVIEW RETURNED	05-Feb-2018

GENERAL COMMENTS	This is an interesting systematic review about the association of political factors with population health outcomes. It is a massive achievement to assess risk of bias, do data extraction and synthesise evidence for 176 studies. The authors also provided a reference list of excluded studies with reasons (full-text screening); it is rare for this to happen even though it is recommended. Nonetheless, the manuscript could benefit from some clarifications. Some of the key issues refer to the provision of further information about 1) the methods used, and 2) both the included studies and the overall results so the readers can make better sense of the findings. Please see further information below: 1. Page 6, lines 3-7: Could the authors clarify what is meant by "proportionate independent second review" for each stage (i.e. study selection, quality assessment and data extraction)? Furthermore, how did the authors solve any disagreements? These issues help us to understand any potential reviewer bias. 2. Page 6, inclusion criteria: Is it possible that the authors have missed many studies by not including grey literature other than scholarly book chapters? Would OECD reports have data, for example? 3. Page 7, first paragraph: Could the authors provide a definition for each of the four political themes? If these were the same definitions from the 2010 review and the exact same framework was adopted, this could be mentioned in the text. 4. Page 10, first paragraph: was successful implementation of effective health policies considered to be a population outcome? Wouldn't this be a system level outcome? Could the authors explain their rationale for including it? 5. Page 26, Table 1: It would be quite useful to see the full search strategy, including appropriate syntaxes, wildcards, truncations and information on whether subject headings (such as MeSH terms) were searched, and on which fields were searched (e.g. title, abstract, keywords). This information helps us to understand how comprehensive the search was, and whether potentially eligible

	articles may have been missed. For example, the search term "globalization" will not find papers that used the British spelling, unless wildcards or subject headings were also used. In this case, a paper with the term "globalisation" would only be found if it also included any of the other search terms. Likewise, publications using the term "democratic" instead of "democracy" may also be relevant, but they will not be identified if neither truncation (e.g. democra*) nor relevant subject headings were searched. Since the team searched several databases, a full search strategy with syntaxes for one database (e.g. Medline) would be sufficient to clarify all these issues. The provision of this search strategy will also help other researchers who may wish to update the review in the future. 6. Pages 27-54, tables 2-5: As many studies were eligible for inclusion, these tables are too large to fit in the main manuscript. Consider having these tables as supplementary data, and synthesising the key findings in one or two smaller tables in the manuscript, perhaps similar to the original review (Muntaner et al 2011), or in any other way the authors see appropriate. Available guidance on narrative synthesis may also be helpful to generate ideas (e.g. Popay et al. Guidance on the Conduct of Narrative Synthesis in Systematic Reviews. A Product from the ESRC Methods Programme. ESRC Methods Programme, 2006). 7. Pages 27-54, tables 2-5: Consider adding the study design, in addition to "individual" and "ecological" as these are very broad definitions. Before and after analyses have different characteristics and potential biases from time series analysis, different types of multivariate analyses generate different questions. Cross-sectional and longitudinal studies also have different strengths and limitations. This additional information will help the reader make sense of the findings, and of the results from the risk of bias assessments. Furthermore, consider adding the aim of each included study to the tables. This also helps the reader to make sense of the included studies and overall findings. 8. Results (overall): the results are repetitive at times, as there is always a differentiation between papers included in the 2010 review and in this review. Consider whether this comparison between the two reviews could be added elsewhere (perhaps the discussion section would be sufficient to highlight this). Furthermore, the results draw the readers' attention to the comparison of findings between reviews, instead of focusing on how all findings tell a story together. Unless the authors wish to focus on the comparison, simplifying the results would make the results section easier to read, and more cohesive. 9. The manuscript refers to the UK context often, and this is not always made clear. It may be worth specifying when this is the case. For example, readers from other countries may not know what NICE is (page 14, line 31). NICE is also an acronym. 10. Are the addition of a risk of bias assessment and further searches sufficient to make this a confirmatory systematic review (as opposed to a "narrative sketch" as proposed in the initial review)? I am a bit concerned about the argument that results support causal inference, considering the range of comparisons made across different countries, using a wide range of research designs, approaching different populations (i.e. adults, children) and including so many different outcomes. Each outcome is likely to have different causal pathways, different confounders, and different covariates, which are not fully captured by ecological studies (and even less captured by surveys and other tools). Hence, it may be more appropriate to have a more cautious approach (even if some of the criteria described by Bradford Hill apply). Due to heterogeneity in

	terms of comparisons, populations, health systems, social structures and outcomes, causal inferences are more difficult to establish. The described outcomes (e.g. self-reported health or depression) are also likely to have multiple causes. Review results are still relevant even if a causal relationship cannot be confirmed. 11. Furthermore, low risk of bias of individual studies does not mean that there is no risk of bias across studies, especially when grouping so many heterogeneous ones. An overall risk of bias rating may also result in unintentional masking of key sources of bias. For example, the overwhelming majority of studies seem to have had moderate risk of bias for selection bias (group equivalence) and chance (sample size rationale). Results were also quite mixed for reporting bias (conflict of interest). Since other sources of bias were less common, the overall scores do not highlight these key issues. This is a limitation that could be acknowledged in the review. 12. Finally, some sub-group analyses may help to interpret the results. I cannot imagine how challenging it was to synthesise so many findings, but at times it seems that the results are not telling the whole story. For example, I can see from Tables 2-5 that the same outcome (e.g. self-rated heath) is shown as a positive, negative or inconclusive across all four political factors (there is a similar issue with life expectancy, limiting long-standing illness and mortality). By simply grouping outcomes in the results section as positive, negative or inconclusive, we are not made aware that there may be conflicting evidence across studies for a specific outcome. This could also be acknowledged as a limitation if the issue is not investigated further.

VERSION 1 – AUTHOR RESPONSE

Comments from Reviewer 1 (David RM Smith, Public Health England)

1. Comment: The authors present a worthwhile update to an existing meta-analysis of the influence of political factors on population health outcomes. Although its conclusions are mostly similar, this update has the benefits of more than doubling the number of included studies and assessing the risk of bias across all included studies.

Response: We thank the reviewer for his appreciation of our work

2. Comment: However, there are important problems with the way the study is presented.

Response: We have sought to address these

3. Comment: The paper would benefit from a clear, neutral introduction to the relationship between political systems and healthcare outcomes, as opposed to the current (rather jumbled) introduction to historical and ideological links.

Response: We have sought to provide a clearer structure to the introduction.

4. Comment: Further, it is obvious that the authors believe that left-wing policies improve health, and are worried about the health consequences of burgeoning right-wing policies and events. They later show that there is evidence-based precedence for these beliefs, but this is a fundamental outcome of the meta-analysis. Introducing this as a foregone conclusion in the introduction undermines the authors' objectivity and questions the need for this study in the first place.

Response: No intention of bias. We have substantially revised the introduction to improve argument structure and phrasing.

5. Comment: Second, the importance of economic development and political history on the generalizability of these results is unclear. To what extent did included studies control for relative social wealth/development? For example, Cereseto & Waitzkin (1986) conclude that leftist policies have positive effects on health within strata of economic development, but that rightist policies may be linked to greater economic development. This is of particular concern with respect to the variable 'political tradition', in which rich/OECD countries are overwhelmingly (but unsurprisingly) over-

represented. But what about elsewhere? The capitalisation of China has entailed enormous economic growth – have healthcare outcomes there not improved substantially as a result? Conversely, the pink tide has recently wrought ruin on several once thriving South American economies – have healthcare systems there not deteriorated? Of course the authors are only working with the evidence they could find, but this kind of global context is important to discuss, especially since the finding that left-wing political traditions always entail better health outcomes is largely based on the recent history of social democracies in the wealthy West.
Response: We have conducted some scenario analyses, and one of them looks at studies that take economic factors into consideration. We have also looked at results according to the type of countries included.
6. Comment: The authors emphasise an internationally comparative approach, but (i) less developed and non-Western nations are underrepresented, and (ii) the narrative focus is clearly on the UK. Perhaps this can be acknowledged, and more global narrative examples could be used, and not just those relevant to Scotland/UK?
Response: We thank the reviewer for this useful insight. We have revised the manuscript to have a stronger international focus and less discussion on Scotland and the UK. We have moved comparison of England and Scotland in the context of devolution to a supplementary data file for readers who are interested in it
7. Comment: The tables assessing risk bias are unwieldy. Offering the colour key in the caption and repeating column headings on each page would help the reader.
Response: We have made this revision as suggested – the colour key is now at the top as well as the bottom, and column headings are now shown at the top of each page.
8. Comment: Table 3 serial number 3: Data are from 1960-1994, not 1960-1964
Response: We have corrected this error, in the table which is now in supplementary file 5.
Comments from Reviewer 2 (Natalia Calanzani, University of Edinburgh)
1. Comment: This is an interesting systematic review about the association of political factors with population health outcomes.
Response: We thank the reviewer for her positive comment about our work
2. Comment: It is a massive achievement to assess risk of bias, do data extraction and synthesise evidence for 176 studies.
Response: We thank the reviewer for her appreciation of the scale of our work. It was certainly a major undertaking.
3. Comment: The authors also provided a reference list of excluded studies with reasons (full-text screening); it is rare for this to happen even though it is recommended.
Response: We thank the reviewer for her appreciation of the thoroughness of our reporting. We think it is important to list studies excluded at full-text stage. In response to other comments about unwieldy results presentation, we have now moved the list of included studies from the reference list to a supplementary file, since there are so many.
4. Comment: Nonetheless, the manuscript could benefit from some clarifications.
Response: We have sought to address these
5. Comment: Some of the key issues refer to the provision of further information about 1) the methods used, and 2) both the included studies and the overall results so the readers can make better sense of the findings. Please see further information below:
Response: We have sought to address these issues and outline our responses below
6. Comment: Page 6, lines 3-7: Could the authors clarify what is meant by “proportionate independent second review” for each stage (i.e. study selection, quality assessment and data extraction)? Furthermore, how did the authors solve any disagreements? These issues help us to understand any potential reviewer bias
Response: We have added text to state to the design sub-section of the methods section: “whereby this author independently appraised 20% of records for each stage. There were few disagreements, and where there arose, they were resolved by discussion.”
7. Comment: Page 6, inclusion criteria: Is it possible that the authors have missed many studies by

not including grey literature other than scholarly book chapters? Would OECD reports have data, for example?
Response: As we acknowledge, one of the limitations of performing an update review rather than a new review was that our inclusion criteria had to be comparable with the 2010 review. This does mean that such reports are ineligible. We have made sure this is clear in the review text by adding the following to the limitations section of the discussion: “Moreover, conducting an update required us to maintain consistency with the 2010 review in terms of inclusion criteria, and precluded us from considering a wider range of grey literature sources, such as Organisation for Economic Co-operation and Development (OECD) reports, which may have relevant data.”
8. Comment: Page 7, first paragraph: Could the authors provide a definition for each of the four political themes? If these were the same definitions from the 2010 review and the exact same framework was adopted, this could be mentioned in the text.
Response: We have added text to the inclusion criteria sub-section of the methods accordingly: “These political features were defined exactly following Muntaner et al”
9. Comment: Page 10, first paragraph: was successful implementation of effective health policies considered to be a population outcome? Wouldn’t this be a system level outcome? Could the authors explain their rationale for including it?
Response: This clearly has population level and system level aspects. We consider the variable in question meets the criteria. This was a record that was subjected to independent second review (as part of the 20% check) and no disagreements arose, so we are satisfied with the inclusion of this study. Also, as this study does not have a positive result, including it errs on the side of caution.
10. Comment: Page 26, Table 1: It would be quite useful to see the full search strategy, including appropriate syntaxes, wildcards, truncations and information on whether subject headings (such as MeSH terms) were searched, and on which fields were searched (e.g. title, abstract, keywords). This information helps us to understand how comprehensive the search was, and whether potentially eligible articles may have been missed. For example, the search term “globalization” will not find papers that used the British spelling, unless wildcards or subject headings were also used. In this case, a paper with the term “globalisation” would only be found if it also included any of the other search terms. Likewise, publications using the term “democratic” instead of “democracy” may also be relevant, but they will not be identified if neither truncation (e.g. democra*) nor relevant subject headings were searched. Since the team searched several databases, a full search strategy with syntaxes for one database (e.g. Medline) would be sufficient to clarify all these issues. The provision of this search strategy will also help other researchers who may wish to update the review in the future.
Response: We have provided an additional supplementary file with the full search strategy for MEDLINE. MeSH terms were used as the primary system. Keywords were also used as a supplementary approach, and key variants were included. The focus on MeSH terms avoids issues such as those you suggest
11. Comment: Pages 27-54, tables 2-5: As many studies were eligible for inclusion, these tables are too large to fit in the main manuscript. Consider having these tables as supplementary data, and synthesising the key findings in one or two smaller tables in the manuscript, perhaps similar to the original review (Muntaner et al 2011), or in any other way the authors see appropriate. Available guidance on narrative synthesis may also be helpful to generate ideas (e.g. Popay et al. Guidance on the Conduct of Narrative Synthesis in Systematic Reviews. A Product from the ESRC Methods Programme. ESRC Methods Programme, 2006).
Response: We have moved these tables to supplementary files. We think that the in text narrative synthesis will now suffice since it has been improved with the removal of the comparison to the previous review, and that summary tables would repeat the in text information and not add anything useful.
12. Comment: Pages 27-54, tables 2-5: Consider adding the study design, in addition to “individual” and “ecological” as these are very broad definitions. Before and after analyses have different characteristics and potential biases from time series analysis, different types of multivariate analyses

generate different questions. Cross-sectional and longitudinal studies also have different strengths and limitations. This additional information will help the reader make sense of the findings, and of the results from the risk of bias assessments. Furthermore, consider adding the aim of each included study to the tables. This also helps the reader to make sense of the included studies and overall findings.
Response: The cross-sectional or longitudinal nature of the studies was already included in the tables since the span of years of data collection is shown. We agreed with the reviewer that more information on study design beyond individual vs ecological would have been useful. However, this information is not regularly reported thoroughly or consistently by the studies, and adding it would have involved a lot of subjective judgement, which we consider unhelpful. We have added this as a limitation to the text. We have added information on whether or not the analysis took economic factors into consideration. We considered that information on study aim would be redundant since exposures and outcomes are listed. Moreover, an important principle of systematic reviewing is to focus on the data available for the review and not to focus on the aims of the original studies.
13. Comment: Results (overall): the results are repetitive at times, as there is always a differentiation between papers included in the 2010 review and in this review. Consider whether this comparison between the two reviews could be added elsewhere (perhaps the discussion section would be sufficient to highlight this). Furthermore, the results draw the readers' attention to the comparison of findings between reviews, instead of focusing on how all findings tell a story together. Unless the authors wish to focus on the comparison, simplifying the results would make the results section easier to read, and more cohesive.
Response: We have changed the results structure accordingly and reduced the focus on comparing results with the previous review. We thank you for your valuable suggestion. We have retained the comparison with the previous review only in the discussion in the sub-section 'Comparison with previous reviews'.
14. Comment: The manuscript refers to the UK context often, and this is not always made clear. It may be worth specifying when this is the case. For example, readers from other countries may not know what NICE is (page 14, line 31). NICE is also an acronym.
Response: We have reduced focus on the UK context in order to appeal more to an international audience. We have moved comparison of England and Scotland to a supplementary file for the benefit of those interested in this.
15. Comment: Are the addition of a risk of bias assessment and further searches sufficient to make this a confirmatory systematic review (as opposed to a "narrative sketch" as proposed in the initial review)?
Response: We have rephrased in order to avoid the word 'confirmatory', and leave that judgement to the readership.
16. Comment: I am a bit concerned about the argument that results support causal inference, considering the range of comparisons made across different countries, using a wide range of research designs, approaching different populations (i.e. adults, children) and including so many different outcomes. Each outcome is likely to have different causal pathways, different confounders, and different covariates, which are not fully captured by ecological studies (and even less captured by surveys and other tools). Hence, it may be more appropriate to have a more cautious approach (even if some of the criteria described by Bradford Hill apply). Due to heterogeneity in terms of comparisons, populations, health systems, social structures and outcomes, causal inferences are more difficult to establish. The described outcomes (e.g. self-reported health or depression) are also likely to have multiple causes. Review results are still relevant even if a causal relationship cannot be confirmed.
Response: We have removed the comments on causality from the summary of findings in the discussion and added discussion of the issue you mention above to the discussion. We have said: "Causative inference can be made more complicated by different causal pathways, different confounders, and different covariates"
17. Comment: Furthermore, low risk of bias of individual studies does not mean that there is no risk of bias across studies, especially when grouping so many heterogeneous ones. An overall risk of bias

rating may also result in unintentional masking of key sources of bias. For example, the overwhelming majority of studies seem to have had moderate risk of bias for selection bias (group equivalence) and chance (sample size rationale). Results were also quite mixed for reporting bias (conflict of interest). Since other sources of bias were less common, the overall scores do not highlight these key issues. This is a limitation that could be acknowledged in the review.
Response: We have added a sub-section to the discussion entitled 'Risk of bias assessment', in which we discuss these issues.
18. Comment: Finally, some sub-group analyses may help to interpret the results. I cannot imagine how challenging it was to synthesise so many findings, but at times it seems that the results are not telling the whole story. For example, I can see from Tables 2-5 that the same outcome (e.g. self-rated heath) is shown as a positive, negative or inconclusive across all four political factors (there is a similar issue with life expectancy, limiting long-standing illness and mortality). By simply grouping outcomes in the results section as positive, negative or inconclusive, we are not made aware that there may be conflicting evidence across studies for a specific outcome. This could also be acknowledged as a limitation if the issue is not investigated further.
Response: We have performed some scenario analyses (could also be called sub-group analyses), which should help clarify matters relating to the results, to the extent that we can given the size of the review (larger reviews are by necessity more high level) and the resources available to us – indeed, our project director was going to be particularly tied up with another project from 5 March for a couple of weeks, so we had to submit revisions early.
Scenario analyses:
1. Studies that take economic factors into consideration, for each of the four themes except globalisation
2. Studies that include developing countries, for each of the four themes
3. Studies looking at general health or quality of life, for each of the four themes
4. Studies using a welfare regime classification scheme, for the welfare state theme
5. Studies using a political tradition classification scheme, for the political tradition theme

VERSION 2 – REVIEW

REVIEWER	David Smith Public Health England, UK
REVIEW RETURNED	04-Apr-2018

GENERAL COMMENTS	The authors have done a thorough job addressing concerns of the reviewers. In particular, inclusion of scenario analyses has offered a deeper level of insight into study findings and has accounted for obvious gaps (e.g., over-representation of Western countries, lack of consideration of economic factors). However, could the authors please include a paragraph discussing these sub-analyses, i.e. how studies accounted for these factors and what was found? In particular, it is not clear how studies would have accounted for economic factors. It is also unclear if any particular developing countries dominated when more developed countries were excluded. Further, a point of clarification: did the scenario analysis of developing countries exclude developed countries? Why or why not? I also appreciate that the authors have toned down their editorial slant, and study discussion and conclusions are now more cautious and balanced. However, more work is needed to ensure the introduction is clear, concise and objective. I encourage the authors to ensure they write precisely what they mean or what was reported in a cited study. Here are three examples from the introduction

	where this was not achieved: (1) “with regard to social determinants of health, there is no formal evidence-based process and ideology dominates evidence” Is this to suggest that countries do not have formal evidence-based processes to evaluate programs to improve social determinants of health, such as education and food security (untrue)? Or perhaps that social determinants of health are not considered in evaluation of which public health policies to adopt (also untrue)? Something else? Regardless, the authors should remember that here they have cited an editorial describing a few recent examples of tension between policy and evidence in the UK -- nothing more and nothing less. (2) “more equal societies have been shown to be more successful” What is meant by “successful”? Success can be measured across innumerable dimensions. (3) “parameters that are static within a given country” – Aren’t the parameters in this review (e.g., political tradition, welfare state investment) also dynamic within countries over time and space? I caution the authors not to undermine their credibility by describing their knowledge of the field as coming from public debates on TV and in newspapers and by emphasizing the non-academic position of one author. At the end of the methods, there is considerable redundancy in describing the role of patients in this study design. I appreciate that the editor asked for these points to be included, but I recommend streamlining the key points and dropping the point-response format. The authors’ conclusions that public health professionals should advocate more is contentious and is not directly supported by the study’s findings. A competing argument is that the academic’s primary role is to elucidate truth, and the clinician’s is to improve the health of their patients. There may well be room for advocacy when professionals’ subject-matter expertise intersects with policy (as highlighted in cited examples), but political motivations among researchers/clinicians can be just as blinding as industry lobbyism, undermining objectivity. This can lead to bad science/medicine and the subsequent erosion of public faith in science/medicine, with potentially far-reaching negative public health consequences. Minor comments: A reference to Muntaner et al. is used to define the political exposure variables in this manuscript. Considering the exposure variables are absolutely central to this work, I encourage the authors to define these terms clearly in the main text. Typo: “There were few disagreements, and where there arose, they were resolved by discussion” Typo (unclosed parentheses): “(MEDLINE, AMED…” Missing N: For welfare state exposure, among countries using a welfare regime classification, n = ? found a positive association?

REVIEWER	Natalia Calanzani University of Edinburgh
REVIEW RETURNED	08-Apr-2018

GENERAL COMMENTS	Thank you to the authors for replying to all my comments; the methods section is clearer and the conclusions more cautious regarding causality. The added text on risk of bias is informative. I have added a few comments below: Comments regarding my initial feedback 1. Item 16 (about assumptions regarding causality): the changes made look good. Consider replacing the word "predictors" (page 12, summary of findings, row 49-50) with a more neutral word (as we are talking about associations). 2. Item 18 (about sub-group analyses): The added analyses are interesting and do help us make sense of the results, but I feel that the issue regarding conflicting evidence for a single outcome has not been addressed with only two outcomes (general health and quality of life). I do understand that this is a high-level review, but if it may lead to misleading conclusions then it is important that something is done about it. Perhaps there could be a table with (more than two) groupings of health outcomes shown in columns, and different political factors shown in rows. The table could be populated with "positives", "negatives" and "inconclusives"). For example, the infant mortality column would have 4 positives, 3 negatives, and 3 inconclusives for the globalisation row. The example also illustrates what I mean by conflicting evidence. Results are quite mixed for infant mortality, but it may be that they are much clearer for other outcomes (for example, for obesity the results seem to be solely negative for globalisation). I also believe that these findings may be more useful in terms of informing readers, discussing causes for the identified associations and making recommendations for policy (compared to reporting the overall proportion of positives and negatives for each political factor). Additional issue: 3. There is an empty column in the risk of bias table that the authors may wish to remove

VERSION 2 – AUTHOR RESPONSE

Reviewer Calanzani	
Comment	**Response**
Consider replacing the word "predictors" (page 12, summary of findings, row 49-50) with a more neutral word (as we are talking about associations).	We have used 'related to' instead
The added analyses are interesting and do help us make sense of the results, but I feel that the issue regarding conflicting evidence for a single outcome has not been addressed with only two outcomes (general health and quality of life). I do understand that this is a high-level review, but if it may lead to misleading conclusions then it is important that something is done about it. Perhaps there could be a table with (more than two) groupings of health	This is provided as Table 3. An additional section is provided at the end of the results section, entitled 'Health outcomes'.

outcomes shown in columns, and different political factors shown in rows. The table could be populated with “positives”, “negatives” and “inconclusives”). For example, the infant mortality column would have 4 positives, 3 negatives, and 3 inconclusives for the globalisation row. The example also illustrates what I mean by conflicting evidence. Results are quite mixed for infant mortality, but it may be that they are much clearer for other outcomes (for example, for obesity the results seem to be solely negative for globalisation). I also believe that these findings may be more useful in terms of informing readers, discussing causes for the identified associations and making recommendations for policy (compared to reporting the overall proportion of positives and negatives for each political factor)	
There is an empty column in the risk of bias table that the authors may wish to remove	Removed
Reviewer Smith	
Comment	**Response**
The authors have done a thorough job addressing concerns of the reviewers. In particular, inclusion of scenario analyses has offered a deeper level of insight into study findings and has accounted for obvious gaps (e.g., over-representation of Western countries, lack of consideration of economic factors). However, could the authors please include a paragraph discussing these sub-analyses, i.e. how studies accounted for these factors and what was found? In particular, it is not clear how studies would have accounted for economic factors. It is also unclear if any particular developing countries dominated when more developed countries were excluded. Further, a point of clarification: did the scenario analysis of developing countries exclude developed countries? Why or why not?	We have separated the text about scenario analyses in the Discussion from the Summary of findings section, and created a separate section for this. In this new sub-section entitled ‘Scenario analyses’, we provide further information to address your request, while also seeking an appropriate balance of section lengths. We have discussed how economic factors were taken into consider. We have considered the country profile among developing countries. We have discussed about whether or not the developing country analysis excluded developed countries.
“with regard to social determinants of health, there is no formal evidence-based process and ideology dominates evidence” Is this to suggest that countries do not have formal evidence-based processes to evaluate programs to improve social determinants of health, such as education and food security (untrue)? Or perhaps that social determinants of health are not considered in evaluation of which public health policies to adopt (also untrue)? Something else? Regardless, the authors should remember that here they have cited an editorial describing a few recent examples of tension between policy and evidence in the UK -- nothing more and nothing less.	We have changed the argument here to make it clearer and more relevant to the focus of the paper. We have dropped the reference to the Bambra editorial and focused on the Ferrie review in the International Journal of Epidemiology, which provides a wider range of examples from different countries, as to how ideology and personal interests can exert substantial influences on policy-making relevant to health.
“more equal societies have been shown to be more successful” What is meant by “successful”? Success can be measured across innumerable dimensions.	We have expanded the text here to list examples of the outcomes studied in The Spirit Level by Wilkinson and Pickett
“parameters that are static within a given country” – Aren’t the parameters in this review (e.g., political tradition, welfare state investment) also dynamic	This was an argument that Muntaner et al used in their article. However, on closer consideration, we agree that it may be inaccurate. We also don’t

within countries over time and space?	consider that it adds much, so this phrase has been removed. The key point here is about 'transcend the particularities of individual countries'.
I caution the authors not to undermine their credibility by describing their knowledge of the field as coming from public debates on TV and in newspapers and by emphasizing the non-academic position of one author	This relates to the public and patient involvement (PPI) section, which we had very little time to write for Revision 1 in response to an editorial office request. We have now rewritten this section, and considered your suggestions in this regard.
At the end of the methods, there is considerable redundancy in describing the role of patients in this study design. I appreciate that the editor asked for these points to be included, but I recommend streamlining the key points and dropping the point-response format.	We have revised and restructured the PPI section, and hope that it is also acceptable to the editorial office
A reference to Muntaner et al. is used to define the political exposure variables in this manuscript. Considering the exposure variables are absolutely central to this work, I encourage the authors to define these terms clearly in the main text.	We have created a Table 2 to provide the requested information within the main manuscript
Typo: "There were few disagreements, and where there arose, they were resolved by discussion"	Corrected
Typo (unclosed parentheses): "(MEDLINE, AMED…"	Corrected
Missing N: For welfare state exposure, among countries using a welfare regime classification, n = ? found a positive association?	Corrected
The authors' conclusions that public health professionals should advocate more is contentious and is not directly supported by the study's findings. A competing argument is that the academic's primary role is to elucidate truth, and the clinician's is to improve the health of their patients. There may well be room for advocacy when professionals' subject-matter expertise intersects with policy (as highlighted in cited examples), but political motivations among researchers/clinicians can be just as blinding as industry lobbyism, undermining objectivity. This can lead to bad science/medicine and the subsequent erosion of public faith in science/medicine, with potentially far-reaching negative public health consequences.	This relates to the 'recommendations for research and academic practice' section. We consider that the content in the 'Implications for policy and practice' section is appropriate, and indeed it does not include the points you refer to about advocacy. We have decided to limit the 'recommendations for research and academic practice' section to become 'Recommendations for future research' and revised content accordingly. Additionally, prior to this, we insert a short section reflecting on published views on the role of academics with relation to the evidence base presented in the manuscript. This is no longer a recommendations section, but discusses published views objectively. We have made a change to the abstract and concluding sentences of the manuscript accordingly.