PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Cohort Profile: HealthWise Wales. A research register and population health data platform with linkage to National Health Service datasets in Wales
AUTHORS	Hurt, Lisa; Ashfield-Watt, Pauline; Townson, Julia; Heslop, Luke; Copeland, Lauren; Atkinson, Mark; Horton, Jeffrey; Paranjothy, Shantini

VERSION 1 – REVIEW

REVIEWER	Cheryl Battersby Imperial College United Kingdom
REVIEW RETURNED	12-Jun-2019

GENERAL COMMENTS	This is a well written manuscript describing the methods of creating a research register and platform. This is a descriptive methodological paper with no analyses of actual data or hypothesis testing. I have several recommendations : 1) In abstract- although the purpose of the research register has been made clear, it is not clear what the aim of the manuscript is. Please state clearly whether it is a) to describe the methods of setting it up or b) to compare the cohort profiles with what is held on data base with general population . If b) then the abstract findings should state how different the cohort is to the general population Paricipants: 21,779 alive vs 43,826 - does this mean a large proportion have now died? This would indicate a very elderly population Main paper: Introduction: Again, the end of introduction could be more specific about the aims of this paper. Is it a descriptive methodological paper or are you setting out to compare cohort versus general. Page 6 lines 23-24 are very broad and relate to the research register not this paper. E.g. why did the authors decide to submit this paper now prior to recruiting the final 50,000 patients? Is it because they want the opportunity to devise methods to inform ongoing recruitment to ensure the population is representative of Wales ? Characteristics or participants :98% of the population in Wales is white- this should be made more obvious in the written results section when ethnicity is mentioned, similar to how sex is. Discussion: Men are very under-represented (28%) and the ages are very specific , and diff to general population. It would be good to mention the implications of this if recruitment does not solve this issue- lack of validity, lack of generalisation etc. The problem with consent based cohort studies is selection and

	recruitment bias. It would also be good to provide more information on how they intend to improve recruitment so we have more men and younger generation. Table and figures- Rather than table 3, I wonder whether the authors have thought of including more details of the 15 studies which recruited 43,826 participants? Overall, this paper is very clear and well written. I would just recommend the authors to be more specific about the aims of this manuscript. i.e. descriptive methodological paper, if that is the intention.

REVIEWER	Jinfeng Zhao The University of Auckland
REVIEW RETURNED	18-Jun-2019

GENERAL COMMENTS	This is an interesting paper on a research register and population health data platform with linkage to National Health Service datasets in Wales. It presents a potentially efficient and cost effective way to recruit participants for population-based research. Congratulations! I would like to provide some minor suggestions below: • While I understand that the structure of a Cohort Profile paper is different from common research paper, the second half of the Collaboration section and the Further details section read like advertisements and instructions for potential users. This information could be pointed to via a hyperlink on your website. • Please provide a clear explanation of what 'research ready' means? • It would be useful to provide a comparison to approaches that use similar platforms (you mentioned the SHARE platform, but didn't discuss it).

VERSION 1 – AUTHOR RESPONSE

Reviewer 1		
Comment		**Response**
5	This is a well written manuscript describing the methods of creating a research register and platform. This is a descriptive methodological paper with no analyses of actual data or hypothesis testing. Overall, this paper is very clear and well written. I would just recommend the authors to be more specific about the aims of this manuscript. i.e. descriptive methodological paper, if that is the intention.	We thank the reviewer for their positive comments.
6	In abstract- although the purpose of the research register has been made clear, it is not clear what the aim of the manuscript is. Please	We have added a sentence to the Abstract (page 2 line 28) and to the final paragraph of the introduction (page 6 line 97), describing the aim of the paper.

	state clearly whether it is a) to describe the methods of setting it up or b) to compare the cohort profiles with what is held on data base with general population . If b) then the abstract findings should state how different the cohort is to the general population… Again, the end of introduction could be more specific about the aims of this paper. Is it a descriptive methodological paper or are you setting out to compare cohort versus general.	
7	Page 6 lines 23-24 are very broad and relate to the research register not this paper. E.g. why did the authors decide to submit this paper now prior to recruiting the final 50,000 patients? Is it because they want the opportunity to devise methods to inform ongoing recruitment to ensure the population is representative of Wales ?	This paper describes the entire HealthWise Wales project. We have submitted the paper now because the project has well-established and tested mechanisms for collecting data, for data linkage to routinely-collected healthcare records, and for researchers to apply to use the platform for their own research. The aim of the paper therefore is to inform researchers of its' existence as a potential resource for their projects (similar to the SHARE Cohort Profile previously published in BMJ Open: McKinstry B et al. Cohort profile: the Scottish Research register SHARE. A register of people interested in research participation linked to NHS data sets. BMJ Open. 2017; 7(2): e013351).
8	Participants: 21,779 alive vs 43,826 - does this mean a large proportion have now died? This would indicate a very elderly population	The registered population is 21,779 individuals. Every individual in the register is given the opportunity to take part in all of the research studies that use the platform for recruitment, as long as they fit the eligibility criteria. Many people have contributed to more than one study. Therefore, the number of people who have contributed to individual studies is much higher than the total number of individuals are registered. We have clarified this by adding a sentence to page 14 line 296 (new text in italics): In total, HWW has facilitated the recruitment of 43,826 participants to 15 different studies to date, *with many of the 21,779 registrants taking part in multiple studies.*
9	Characteristics or participants :98% of the population in Wales is white- this should be made more obvious in the written results section when ethnicity is mentioned, similar to how sex is.	We have added a percentage to the sentence where we talk about the percentage of the study population and the general population who are from non-white ethnic groups (page 13 line 274)
10	Discussion: Men are very under-represented (28%) and the ages are very specific , and diff to general population. It would be good to mention the implications of this if recruitment does not solve this issue- lack of validity, lack of generalisation etc. The problem with consent based cohort studies is selection and recruitment bias.	Many large cohort studies and/or registers such as this one tend to recruit samples that are not representative of the general population. For example, it is known that the SHARE register in Scotland and the UK Biobank study have recruited more women than men, and fewer individuals from deprived areas. The aim of a register such as this is to recruit enough individuals with different characteristics so that appropriate samples from the recruited population can be selected on a study-by-study basis. We explain this in the paper on page 15 (line 333): *The aim is to achieve a study sample that closely models the population of Wales, with sufficient numbers in socio-*

	It would also be good to provide more information on how they intend to improve recruitment so we have more men and younger generation.	*demographic subgroups to allow for the selection of populations for research from those groups. For example, the cohort currently includes 5,000 men, providing a substantial sample size that will be adequate for some analyses.* We agree, however, that we now need to develop recruitment mechanisms that specifically target the under-represented groups, and that it is useful to provide information on this in the paper. We have added a description of the current plans for this on page 15 (line 321) that leads into the sentences noted above.
11	Table and figures- Rather than table 3, I wonder whether the authors have thought of including more details of the 15 studies which recruited 43,826 participants?	The HealthWise Wales team are not the principal investigators of the studies that have used the platform for recruitment. All of the studies listed will publish their results individually once they have been completed. The information provided is given in a similar format to the Cohort Profile paper which describes the SHARE register. In addition, as most of these studies are ongoing, we have provided enough detail to give a flavour of the studies that have used the platform to date without jeopardising the publication of the studies themselves. We have now added information on a paper that has been published using data from participants recruited through HWW, and a sentence describing the location of information on publications relating to studies that have used HWW on page 14 (line 297 onwards).
Reviewer 2		
Comment		**Response**
12	This is an interesting paper on a research register and population health data platform with linkage to National Health Service datasets in Wales. It presents a potentially efficient and cost effective way to recruit participants for population-based research. Congratulations!	We thank the reviewer for their positive comments.
13	While I understand that the structure of a Cohort Profile paper is different from common research paper, the second half of the Collaboration section and the Further details section read like advertisements and instructions for potential users. This information could be pointed to via a hyperlink on your website.	We agree that the amount of information in this section can be reduced, and have removed several lines of text on page 18 (lines 379-387).
14	Please provide a clear explanation of what 'research ready' means?	We have removed this term from the abstract, and replace it with "*a cohort of individuals who have consented to be informed about research opportunities*" (page 2, line 25). We have also changed the sentence where this term first appears (on page 6, line 89) to read: *Establishing a cohort of individuals who have consented to be contacted with information on research studies that they may wish to contribute to (so-called "research-ready" individuals)*
15	It would be useful to provide a comparison to approaches that use	We agree that adding information to compare HWW with similar platforms is a useful addition to the paper. We have

	similar platforms (you mentioned the SHARE platform, but didn't discuss it).	included this information in the "strengths and limitations" section on page 16 (line 339 onwards).

VERSION 2 – REVIEW

REVIEWER	Cheryl Battersby Imperial College , United Kingdom
REVIEW RETURNED	01-Sep-2019

GENERAL COMMENTS	I am happy with the authors' revisions in response to my original comments. I think this is an important resource that should be widely publicised among researchers and others who will benefit.