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)	UNDERSTANDING REPEATED NON-ATTENDANCE IN HEALTH SERVICES - PILOT ANALYSIS OF ADMINISTRATIVE DATA AND FULL STUDY PROTOCOL FOR A NATIONAL RETROSPECTIVE COHORT
AUTHORS	Williamson, Andrea; Ellis, David; Wilson, Philip; McQueenie, Ross; McConnachie, Alex

VERSION 1 - REVIEW

REVIEWER	Dr Erica Cook University of Bedfordshire, United Kingdom
REVIEW RETURNED	27-Sep-2016

GENERAL COMMENTS	This is really important research which will have a huge impact on how we can strive to reduce health inequalities. I have outlined a few areas in which I think this paper could be improved. The focus group contains 5 GPs which were purposefully selected to represent a range of perspectives. However, it was unclear what perspectives they were trying to cover. I am assuming from the additional file that it is from areas of urban/rurality deprivation/affluence mix but this is not clear from the onset. How is deprivation defined here? How is semi-rural defined? I am also interested here about the GP practice size that these represent – not sure if this information is available and the cohort distributions they serve i.e. BME population? Older population etc. It was not clear why a focus group was conducted i.e. in favour of individual interviews? Did these GP's know each other? How did the dynamics work? How did you manage to get 5 GP's together at the same time etc. The results of the focus group are then presented with no information of how the data were analysed. There was also no data to support the findings. I did not really get a feel for how the GP's views differed and how important these differences were. For example, there is a discussion around the prevalence of missed appointments in affluent areas and deprived areas but it was not clear how prevalence differed? And why it differed? There is an interesting point made about marginalised patients but who are these? I also think that there are some clear stereotypes and negative perceptions e.g. 'patients do not value free health care'. It is also stated that some practices have a negative view – who? Why? This

	also links into issues of workload and expectations. Overall, I think the data needs to be better organised and presented so that important findings are given the discussion they need. It is worth stating what a community Health Index number is. I was also not sure if all data had a CHI and if not how probabilistic matching would be used. The statistical analysis section is where all the data that will be linked is reported. I think this will be better suited to a dataset section. The statistical analysis can then just be focused on how the data will be analysed.

REVIEWER	Emely Ek Blæhr DEFACTUM, Central Denmark Region Aarhus University Aarhus, Denmark
REVIEW RETURNED	09-Nov-2016

GENERAL COMMENTS	This manuscript is a protocol and pilot study of a very exciting and concerning study with a mixed methods approach in the pilot and a retrospective cohort design in the full study, about determining the relationship between general practice appointment attendance and health care perspectives in a life course approach. The manuscript includes some very good considerations about ethics. I have some minor concerns, detailed below: 1. First of all, I suggest that the author uses a more clear structure in the Methods and analysis section. According to guidelines for pilot studies, the "Aim and Research questions" should be moved up after the introduction, but before the methods section. The study/research questions should not be repeated. They should be in the end of the introduction or in the aims section. Further, methods and results should be more clearly separated in the pilot study (e.g. line 213 page 10) 2. Which protocol study design is used? (E.g. CONSORT, SPIRIT etc.) The structure of this method should be clearer. 3. It would be very useful to have a description of the population at the beginning of the methods section. 4. At line 216-218, page 10 you assume that your final sample provides a similar distribution as the sample from the pilot study and on that behalf you will classify patients based on the number of appointments missed over the last three years. Why is this an appropriate assumption? 5. Do you have any methodical thoughts regarding the rate of attending appointments as a potential predictor of the rate of non-attendance? There may be some conflicting interactions to control for. 6. Please state more clearly, which predictors and factors you will analyze (and in what table you can find them). 7. Please spell abbreviations out the first time they are used (E.g. Page 2 line 30, page 8 line 163). 8. Please write the rest of the sentence (page 13, line 278) 9. Elaborate on the term "Read codes"; please include a reference if possible. 10. According to the BMJ guidelines, include dates of study and data analysis plan 11. Consider if line 44-45 page 3 should be moved to the section

	"Ethics and dissemination" – it is not a part of the method or analysis 12. The categories for number of missed appointments is not mutually exclusive – is the last category >2?

VERSION 1 – AUTHOR RESPONSE

'Reviewer 1
This is really important research which will have a huge impact on how we can strive to reduce health inequalities.'
Thank you- that is positive to hear.

'I have outlined a few areas in which I think this paper could be improved.'
This is helpful feedback about how we have represented the qualitative focus group in this paper. We were uncertain as to how much data to report. We have added in and explained the detail as requested from the comments below. However this must also (in our view) be balanced with this being one focus group of 5 GPs. We do not wish to attract criticism about giving disproportionate weight to these results.

'The focus group contains 5 GPs which were purposefully selected to represent a range of perspectives. However, it was unclear what perspectives they were trying to cover. I am assuming from the additional file that it is from areas of urban/rurality deprivation/affluence mix but this is not clear from the onset. How is deprivation defined here? How is semi-rural defined? I am also interested here about the GP practice size that these represent – not sure if this information is available and the cohort distributions they serve i.e. BME population? Older population etc'.
Additional information is included (page 9).

'It was not clear why a focus group was conducted i.e. in favour of individual interviews? Did these GP's know each other? How did the dynamics work? How did you manage to get 5 GP's together at the same time etc.'
Further explanation is included (page 9).

'The results of the focus group are then presented with no information of how the data were analysed. There was also no data to support the findings. I did not really get a feel for how the GP's views differed and how important these differences were. For example, there is a discussion around the prevalence of missed appointments in affluent areas and deprived areas but it was not clear how prevalence differed? And why it differed? There is an interesting point made about marginalised patients but who are these?'
More information is provided (page 9-10).

'I also think that there are some clear stereotypes and negative perceptions e.g. 'patients do not value free health care'. It is also stated that some practices have a negative view – who? Why? This also links into issues of workload and expectations. Overall, I think the data needs to be better organised and presented so that important findings are given the discussion they need.'
We hope this has been achieved - also taking into account our comment about proportionality above (page 10-11).

'It is worth stating what a community Health Index number is.
I was also not sure if all data had a CHI and if not how probabilistic matching would be used'.
This has been added in (page 17).

'The statistical analysis section is where all the data that will be linked is reported. I think this will be better suited to a dataset section. The statistical analysis can then just be focused on how the data

will be analysed.'
This is helpful feedback too - based on this and the feedback from reviewer 2 and the editor, revisions to the structure of the paper have been made based as closely as possible on the STROBE guidelines (see below).

'Reviewer: 2
This manuscript is a protocol and pilot study of a very exciting and concerning study with a mixed methods approach in the pilot and a retrospective cohort design in the full study, about determining the relationship between general practice appointment attendance and health care perspectives in a life course approach. The manuscript includes some very good considerations about ethics.'
Thank you for your positive feedback.

'I have some minor concerns, detailed below:

First of all, I suggest that the author uses a more clear structure in the Methods and analysis section. According to guidelines for pilot studies, the "Aim and Research questions" should be moved up after the introduction, but before the methods section.'
This has been done (page 7).

'The study/research questions should not be repeated. They should be in the end of the introduction or in the aims section. Further, methods and results should be more clearly separated in the pilot study (e.g. line 213 page 10)'
This has been done and we hope it makes better sense to the reader (pages 8-14).

'Which protocol study design is used? (E.g. CONSORT, SPIRIT etc.) The structure of this method should be clearer. '
Because of the interaction between the results from the pilot study impacting on the protocol design and the design of both we found it challenging to ensure our paper followed a specific study design checklist. However because it is a mixed cross-sectional and cohort study we used the STROBE checklist. A completed checklist is now included in the submission. All comments have been really helpful in improving the structure of the paper- appreciated. We hope they now better reflect all suggestions.

'It would be very useful to have a description of the population at the beginning of the methods section'.
This has been done (page 7).

'At line 216-218, page 10 you assume that your final sample provides a similar distribution as the sample from the pilot study and on that behalf you will classify patients based on the number of appointments missed over the last three years. Why is this an appropriate assumption?'
GP practices were selected so that we were likely to get a representative sample at both the pilot data stage and the final sample. Because our pilot sample was large, it is appropriate to assume that this will scale-up accordingly and we have no reason to believe that it wouldn't. The distribution also suggested categories based on integer numbers of missed appointments per year which will be helpful for policy and clinical stakeholders (page 14).

'Do you have any methodical thoughts regarding the rate of attending appointments as a potential predictor of the rate of non-attendance? There may be some conflicting interactions to control for.'
We control for this using an offset term in our negative binomial model which accounts for number of appointments made (page 20).

'Please state more clearly, which predictors and factors you will analyze (and in what table you can find them)'. A table that summarizes the categories and overall variables for analysis has been included (page 25-26).

'Please spell abbreviations out the first time they are used (E.g. Page 2 line 30, page 8 line 163).'
We have conducted a careful check and amended accordingly.

'Please write the rest of the sentence (page 13, line 278)' Thank you- sentence now completed (page 16).

'Elaborate on the term "Read codes"; please include a reference if possible.' This has been included (page 21).

'According to the BMJ guidelines, include dates of study and data analysis plan' Thanks yes- also flagged by editor and reviewer 1, suggestions incorporated (page 7).

'Consider if line 44-45 page 3 should be moved to the section "Ethics and dissemination" – it is not a part of the method or analysis.'
Agree and on re-reading meaning already contained in dissemination- so deleted (page 3).

'The categories for number of missed appointments is not mutually exclusive – is the last category >2? '
Thank you- this was an overlooked important typo, now corrected and the text has been re-written to make this categorisation clearer for the reader (page 14).

VERSION 2 – REVIEW

REVIEWER	Erica Cook University of Bedfordshire United Kingdom
REVIEW RETURNED	28-Dec-2016

GENERAL COMMENTS	The authors have addressed all comments suggested.

REVIEWER	Emely Ek Blæhr DEFACTUM, Denmark
REVIEW RETURNED	05-Jan-2017

GENERAL COMMENTS	Thank you for your good response. The manuscript has been revised very well and I am especially fond of the table of overall variables (page 25+26). My only (minor)comments is regarding the STROBE checklist where I will recommend to include the table of variables (page 25+26) in the variables section and in the quantitative variables section. Further, one could argue to extend the section on statistical methods.