PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Ventilatory function as a predictor of mortality in lifelong non-smokers: evidence from large British cohort studies
AUTHORS	Gupta, Ramyani; Strachan, David

VERSION 1 - REVIEW

REVIEWER	David Batty UCL, UK
REVIEW RETURNED	30-Dec-2016

GENERAL COMMENTS	bmjopen-2016-015381 - Spirometry and survival in large UK cohorts of lifelong non-smokers The purpose of the paper is never made clear or made persuasive – 'verification' is a poor justification. I think the work has various utilities – FEV1 vs. FVC in relation to mortality; lung function vs. other mortality risk factors; multiple studies etc – but the authors need to be explicit. 'Survival' in the title implies the authors are examining the relation of lung function with death (prognosis) after diagnosis of a chronic disease, which is, largely, not the case. Also, the use of the 'spirometry' term in the title is odd - if this was a paper in which blood pressure was the exposure of interest, one would not use sphygmomanometer in the title. Suggest edits along the following lines: "Lung function in relation to total mortality and cause-specific mortality in lifelong non-smokers: evidence from six prospective cohort studies" I would have liked to have seen the data analysed using a 2-stage individual participant meta-analytical (IPMA) technique in which results are first presented on a study-specific basis, then pooled. In this way, one gets to explore results according to each study. Currently, as I read it, there is no recognition that the studies might produce different results – presumably the control for study year is a way of dealing with secular increases in lung function rather than study variation. IPMA is widely used now and is an approach applied to HSE and SHS cohorts: https://www.ncbi.nlm.nih.gov/pubmed/22849956 (plus other papers) I do not think that having a geographically representative sample (SHS/HSE) is important when exploring disease aeitiology. We've shown that risk factors for CVD reveal near-identical relationships in highly select groups (Whitehall II) and in general population based cohort studies (BRHS): https://www.ncbi.nlm.nih.gov/pubmed/25265141

	I don't see the relevance of having comments about the future utility of UKBB in the strengths and limitations section (pg. 4) Strictly 'association' rather than 'effect'. Ref. 4 (which should be retained) has been updated, albeit using analyses of cohort data as opposed to a case-control study: https://www.ncbi.nlm.nih.gov/pubmed/17119881 I'd like to see more discussion of extant studies (ref 1-7) in the introduction. Give the correlation between the various lung function measures – you state it is strong. Pg. 5 – surely 'data' not 'results' were combined. Pg. 8 – 'statistical analyses' rather than 'modelling of mortality'. Here, was the Cox assumption of proportional hazards supported? Explain why survival analyses were also partitioned at 5 years. This is presumably to explore reverse casualty. Pg. 14 – HSE and SHS have data on baseline respiratory function too, as I recall. Discussion: use of subtitles would help; I'd like to read some text on mechanisms, particularly for the lung function–non-respiratory disease mortality relationships which, if not a new result, is intriguing. Where is the conclusion section?

REVIEWER	Joachim Heinrich LMU, Institute of Occupational, Social, and Environmental Medicine, Germany
REVIEW RETURNED	24-Feb-2017

GENERAL COMMENTS	Bmjopen-2016-015381 The role of lung function on longevity seems to be well established, but its specific prediction on mortality among never smokers is insufficiently clear. The aim of this study is to model the predictivity of spirometric lung function data on all-cause and cause-specific mortality in never smokers. The authors used data from two very large cohorts from the UK, which differ in several methodological aspects. In addition, the reduction of the determinants of mortality to just a single characteristic (spirometric lung function) is challenging. The authors structured the manuscript very well, provided all necessary data and the methodological limitations are discussed in a very transparent way. The appropriate statistical models were also applied. The key findings add to the current knowledge and are supported by a thoughtful discussion. Although the topic is complex, the manuscript is easy to follow. However, there are a few minor comments, which could be used to further improve the manuscript. Minor 1. Although the discussion mentions the two previous papers on the same topic, the results from these previous papers should be added in order to let the readers know what the current paper adds. 2. Abstract: The findings for mortality from respiratory diseases are not mentioned in the abstract at all and even the direction for the cancer mortality is not given.

	3. Article summary: The fourth bullet point is not strongly related to this study and could be removed. 4. P8, line 8: Please present the number of excluded subjects with outlying observations. 5. P11, line 12: This statement does not hold for FVC. 6. P14, section starting with line 18: The role of obesity is not sufficiently studied throughout the manuscript. BMI was not used for GLI adjustment, although BMI might have an impact on lung function testing. Further, BMI is a well-established predictor of mortality. So far, only the HR were additionally adjusted for BMI. The supplementary e-Table 2 could be extended to include results for obese versus non-obese subjects. 7. P15, table 3: The authors correctly restricted the comparisons between the groups with and without a specific condition. However, the reader might (incorrectly) compare the given effect estimates between conditions at entry. Do the authors think that a comparison with an “apparently healthy” group for all conditions at entry might be more informative? This information could alternatively be provided in a table in the supplement? 8. P16, last section starting with line 46: Why are the results shown for all non-respiratory deaths, but not for all-cause deaths, the latter of which is the main outcome of interest? 9. P16, line 56: SBP does not show a similar pattern between the two cohorts. 10. P28, e-Table 2: Have the authors considered discussing the results presented in e.Table 2. Specifically, that current smokers have sometimes a lower mortality risk than former smokers?

REVIEWER	Arnulf Langhammer NTNU, Norwegian University of Science and Technology, Faculty of Medicine and Health Care, Trondheim, Norway
REVIEW RETURNED	04-Mar-2017

GENERAL COMMENTS	This is an interesting study based on a large number of spirometry data from the Health Surveys for England, the Scottish Health Survey and the UK Biobank. The lung function indices are reported as z-scores calculated by the GLI-2012 software. This should be written explicitly and referred as GLI-2012 reference values and GLI-2012 z scores. The manuscript includes many tables with comprehensive data. The text is a bit difficult to follow as comparisons of numbers in the tables does not seem to be tested by significance testing and different patterns are found among never, former and current smokers but not referred correctly in the text. The manuscript had been easier to follow if there were more specific aims, and statistics and results were reported in the same order. Comparisons between different qualities of spirometry and between the measuring sites are of interest, but shouldn’t rather the comparisons with other measures like BMI and BP have been given priority. The unexpected U-shape for BMI and BP compared to linear relation between quartiles of z-scores should be discussed more thoroughly. Comments: Abstract line 22-24: GLI-2012 z-scores are standardized measures not standard deviations. Among causes of death respiratory diseases should also be included. Introduction: The authors write that the HSE and SHS are more representative of the general population than UK Biobank. What was the participation rate in the age group 40-69 years in HSE/SHS? In

	what way did the participants differ between HSE/SHS and UKB? Have any nonparticipation studies been performed? The statistics should be described in more detail: There should be a table reporting descriptive data of the study samples being compared; All, UKB, HSE-SHS, those with non-respiratory mortality, circulatory mortality etc. The GLI z-scores are standardized measures of lung function adjusted for age, sex and height. If there is lack of fit for GLI-2012 in the studied population, the z-scores could vary by these confounders. Is this the rationale for adjusting for age, sex and height in the proportional hazard models? In many tables HR are reported and compared; in the statistics it should be stated whether assumptions were violated or met. The tables report HR from many models; how has the authors tested HR from different models? The legend of table 1 should include by GLI Z scores for FEV1, FVC and FEV1/FVC. According to the footnotes the HRs are expressed per SD decrement in z-score; this is not correct. Page 11 line 11; what is meant by the most direct comparison? I cannot see that the statement: Although the associations of spirometric indices with these earlier deaths were stronger in HSE-SHS than in UKB…. is correct according to data given in e-table 1. Further, how was significance of differences tested? I cannot see that the statement in line 27-31 is in accordance with data in e-table 2. Page 14 line 20-25. The text regarding male never smokers cannot be correct. And slightly stronger association than?? Page 14 line 35-42. Is it correct that the pattern of results was consistent between the age categories in all smoking subgroups?

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author: Reviewer: 1

The purpose of the paper is never made clear or made persuasive – 'verification' is a poor justification. I think the work has various utilities – FEV1 vs. FVC in relation to mortality; lung function vs. other mortality risk factors; multiple studies etc – but the authors need to be explicit.
5. Both the abstract (p3) and the introductory section (p5-6) has been extended to include more detail on the purpose. The paper addresses multiple aims, but we see this as a strength rather than a limitation.

'Survival' in the title implies the authors are examining the relation of lung function with death (prognosis) after diagnosis of a chronic disease, which is, largely, not the case. Also, the use of the 'spirometry' term in the title is odd - if this was a paper in which blood pressure was the exposure of interest, one would not use sphygmomanometer in the title. Suggest edits along the following lines: "Lung function in relation to total mortality and cause-specific mortality in lifelong non-smokers: evidence from six prospective cohort studies"
6. The title has been amended as in point 1 above. We prefer "ventilatory function" to "lung function" because there are other aspects of lung function (eg. gas exchange, ventilation-perfusion imbalance) which are not measured by spirometry. "Spirometry" (rather than "spirometer") is used as a shorthand for measurements of ventilatory function.

I would have liked to have seen the data analysed using a 2-stage individual participant meta-analytical (IPMA) technique in which results are first presented on a study-specific basis, then pooled. In this way, one gets to explore results according to each study. Currently, as I read it, there is no recognition that the studies might produce different results – presumably the control for study year is a

way of dealing with secular increases in lung function rather than study variation. IPMA is widely used now and is an approach applied to HSE and SHS cohorts:
https://www.ncbi.nlm.nih.gov/pubmed/22849956 (plus other papers)
7. We have investigated the 2-stage IPMA approach for combining the results of the five national health surveys (Health Surveys for England and Scottish Health Surveys). The results for all-cause mortality in lifelong non-smokers are shown in annex 1, in the form of forest plots of the study-specific hazard ratios for zFEV1 and zFVC.
These forest plots illustrate that there is no substantial (and certainly no statistically significant) heterogeneity of effect among the five national health surveys, nor between the pooled result for the national health surveys and the corresponding result for UK Biobank. However, UK Biobank alone accounts for almost four-fifths of the weighting in the overall meta-analysis, whereas the individual national health surveys each account for only 2-7% of the overall weight.
In our manuscript, we include all five national health surveys in a single model, adjusting for survey year as a categorical variable, which approximates to a fixed-effect pooling of the study-specific associations between spirometric indices and mortality. (The difference between this approach and a fixed-effect 5-study IPMA is that the effect of covariates is allowed to differ between studies in the IPMA approach.)
We have investigated four alternative approaches to this pooled analysis, for all-cause mortality in never-smokers:
a) Cox regression of the combined surveys with adjustment for survey as a categorical covariate (as presented in the paper)
b) Cox regression of individual surveys followed by fixed-effect meta-analysis (as shown in the upper portion of the forest plots)
c) Poisson regression of the combined surveys with adjustment for survey as a categorical covariate (ie. a fixed intercept for each of the five surveys)
d) Multi-level Poisson regression of the combined surveys including a random intercept for each of the five surveys.
Each of these approaches generates very similar results for zFEV1 and for zFVC in relation to all-cause mortality in lifelong non-smokers, as shown in annex 1, and now mentioned at page 9.
We have therefore retained our original analytical approach (including all 5 national health surveys in a single proportional hazards model, adjusted for survey year) for all results presented in our revised manuscript.
We prefer to analyse UK Biobank separately because:
a) Its covariates are defined differently to the national health surveys (Biobank recruitment centre rather than region and Townsend deprivation score rather than household social class)
b) One of the purposes of the paper is to investigate how closely the Biobank results match those from national health surveys.

I do not think that having a geographically representative sample (SHS/HSE) is important when exploring disease aeitiology. We've shown that risk factors for CVD reveal near-identical relationships in highly select groups (Whitehall II) and in general population based cohort studies (BRHS):
https://www.ncbi.nlm.nih.gov/pubmed/25265141
8. We acknowledge that in the specific comparison to which reviewer 1 refers, the risk factor associations are similar for a London-based occupational cohort and for a geographically dispersed sample recruited through general practices. However, we would argue strongly that "one swallow does not make a summer" and that evidence of generalisability from UK Biobank to more representative samples is a valuable contribution and therefore that seeking such evidence is a valid aim for our paper.

I don't see the relevance of having comments about the future utility of UKBB in the strengths and limitations section (pg. 4)
9. We accept this point. This section (p4) has been rewritten (see point 2 above).

Strictly ‘association’ rather than ‘effect’.
10. We are not sure which sentence this refers to, but have made the change throughout.

Ref. 4 (which should be retained) has been updated, albeit using analyses of cohort data as opposed to a case-control study: https://www.ncbi.nlm.nih.gov/pubmed/17119881
11. This reference has been added in the discussion section (ref 19, pages 19-20).

I’d like to see more discussion of extant studies (ref 1-7) in the introduction.
12. Rather than extend the introduction, we have included comparisons between our study and these papers in the discussion section (p19-20). However, we have included some discussion of reference 7 in the introduction (p5) as the comparison of associations with FEV1 and FVC raised by this paper is one of our aims.

Give the correlation between the various lung function measures – you state it is strong.
13. The inter-correlation of zFEV1, zFVC and zFEVFVC among lifelong non-smokers in HSE-SHS (5 surveys combined) and in Biobank is now mentioned in the methods section (p9).
Correlations: Biobank HSE-SHS (40-69) HSE-SHS (70+)
zFEV v zFVC 0.8753 0.8029 0.8138
zFEV v zFEVFVC 0.3468 0.3609 0.4804
zFVC v zFEVFVC -0.1312 -0.2012 -0.0542
Number of subjects 149,343 6,579 1,429

Pg. 5 – surely ‘data’ not ‘results’ were combined.
14. This has been amended at the start of the methods section (p6).

Pg. 8 – ‘statistical analyses’ rather than ‘modelling of mortality’. Here, was the Cox assumption of proportional hazards supported? Explain why survival analyses were also partitioned at 5 years. This is presumably to explore reverse casualty.
15. We prefer to retain the original subtitle because “statistical analysis” also applies to the preceding section (Adjustment of spirometric measures for gender, age and height). For lifelong non-smokers, the assumption of proportionality of hazards was supported when time-dependence of the hazards ratio was formally tested for statistical significance (see more detail in annex 2, in response to reviewer 3, point 36 below, and page 12 in the paper).
The rationale for presenting all-cause mortality results partitioned at 5 years of follow-up was indeed partly to exclude reverse causality (ie. spirometry compromised by pre-existing conditions which led to early death – although the analyses presented in Table 3 by prior disease history also address this point). However, the main reason was that the minimum length of follow-up in UK Biobank was 4.87 years in each smoking subgroup, so virtually all that cohort had been followed for 5 years or more, allowing a more direct comparison with the national health surveys, all of which had been followed for more than 5 years. (See p9.)

Pg. 14 – HSE and SHS have data on baseline respiratory function too, as I recall.
16. Assuming that this comment refers to baseline “respiratory disease” (rather than “function”), we agree, but chose to present the comparisons in Table 3 only for UK Biobank because of its much larger size and therefore greater power to detect (or exclude) possible effect modification by pre-existing disease.

Discussion: use of subtitles would help; I’d like to read some text on mechanisms, particularly for the lung function–non-respiratory disease mortality relationships which, if not a new result, is intriguing. Where is the conclusion section?
17. We have included subtitles in the discussion section (p18-21), following the editorial guidelines,

and added a brief conclusion (p22). Regarding mechanisms, we have alluded briefly to the possible lifecourse influences on ventilatory function, with reference to the reviewer's own paper on the Whitehall I cohort (see point 11).

Reviewer(s)' Comments to Author: Reviewer 2

The role of lung function on longevity seems to be well established, but its specific prediction on mortality among never smokers is insufficiently clear. The aim of this study is to model the predictivity of spirometric lung function data on all-cause and cause-specific mortality in never smokers. The authors used data from two very large cohorts from the UK, which differ in several methodological aspects. In addition, the reduction of the determinants of mortality to just a single characteristic (spirometric lung function) is challenging. The authors structured the manuscript very well, provided all necessary data and the methodological limitations are discussed in a very transparent way. The appropriate statistical models were also applied. The key findings add to the current knowledge and are supported by a thoughtful discussion. Although the topic is complex, the manuscript is easy to follow. However, there are a few minor comments, which could be used to further improve the manuscript.

Minor
1. Although the discussion mentions the two previous papers on the same topic, the results from these previous papers should be added in order to let the readers know what the current paper adds.
18. This has been added to the discussion at pages 19-20 (see also point 12 above).

2. Abstract: The findings for mortality from respiratory diseases are not mentioned in the abstract at all and even the direction for the cancer mortality is not given.
19. These have been added to the abstract at page 3 (see also point 32 below).

3. Article summary: The fourth bullet point is not strongly related to this study and could be removed.
20. This section has been rewritten at page 4 (see also points 2 & 9 above).

4. P8, line 8: Please present the number of excluded subjects with outlying observations.
21. A comment has been added to the methods section (page 8). The details are shown here:
Excluded outliers: Never smokers Former smokers Current smokers
UK Biobank 303/149646 (0.20%) 298/191925 (0.15%) 74/38588 (0.19%)
HSE-SHS (40-69) 36/6615 (0.54%) 46/9449 (0.49%) 38/6678 (0.57%)
HSE-SHS (70+) 4/1433 (0.28%) 12/3095 (0.39%) 1/910 (0.11%)

5. P11, line 12: This statement does not hold for FVC.
22. This paragraph, now on page 12, has been amended.

6. P14, section starting with line 18: The role of obesity is not sufficiently studied throughout the manuscript. BMI was not used for GLI adjustment, although BMI might have an impact on lung function testing. Further, BMI is a well-established predictor of mortality. So far, only the HR were additionally adjusted for BMI. The supplementary e-Table 2 could be extended to include results for obese versus non-obese subjects.
23. The obese/non-obese comparison has been added to the supplement (now e-Table 3), with a brief discussion of the findings on page 15.

7. P15, table 3: The authors correctly restricted the comparisons between the groups with and without a specific condition. However, the reader might (incorrectly) compare the given effect estimates between conditions at entry. Do the authors think that a comparison with an "apparently healthy" group for all conditions at entry might be more informative? This information could alternatively be

provided in a table in the supplement?
24. Table 3 presents results stratified by specific pre-existing conditions. We feel this is a more powerful and appropriate way to address the issue of reverse causality than to attempt to define an "apparently healthy" group, which would beg questions about what range of pre-existing conditions should be considered to exclude a participant from this group.

8. P16, last section starting with line 46: Why are the results shown for all non-respiratory deaths, but not for all-cause deaths, the latter of which is the main outcome of interest?
25. The presentation of results for non-respiratory mortality in our original submission was designed to be conservative in relation to the association of spirometric indices with mortality, showing that even if respiratory deaths are excluded, there are stronger relationships with zFEV1 and zFVC than with BMI or SBP. However, we do take the point that the main thrust of the rest of the paper is all-cause mortality, so we have revised the figure and the associated data table in the supplement (now e-Table 7) to show results for all-cause mortality and circulatory mortality, rather than non-respiratory mortality and circulatory mortality. The text at pages 17-18 has been amended accordingly.

9. P16, line 56: SBP does not show a similar pattern between the two cohorts.
26. The patterns for non-respiratory mortality (and for all-cause mortality in the revised figure) are similar in relation to SBP in the national health surveys and in UK Biobank. The patterns appear different for circulatory mortality, but the confidence intervals in HSE-SHS are very wide and no firm conclusion can be drawn for this subset. We have amended the statement at the top of page 18 to refer more specifically to all-cause mortality.

10. P28, e-Table 2: Have the authors considered discussing the results presented in e.Table 2. Specifically, that current smokers have sometimes a lower mortality risk than former smokers?
27. The results in former e-Table 2 (now e-Table 3) are hazard ratios (per unit decrement in z-score), not absolute mortality rates. It is perhaps not so remarkable that the gradient of mortality risk across the range of spirometric performance is greater for ex-smokers than the corresponding gradient of risk among current smokers. Former smokers who gave up a long time ago will have better preserved ventilatory function and lower mortality risk than those who ceased smoking more recently. Therefore, among ex-smokers, spirometric indices partly reflect lifetime exposure to active smoking.

Reviewer(s)' Comments to Author: Reviewer 3

This is an interesting study based on a large number of spirometry data from the Health Surveys for England, the Scottish Health Survey and the UK Biobank. The lung function indices are reported as z-scores calculated by the GLI-2012 software. This should be written explicitly and referred as GLI-2012 reference values and GLI-2012 z scores.
28. This change has been made throughout.

The manuscript includes many tables with comprehensive data. The text is a bit difficult to follow as comparisons of numbers in the tables does not seem to be tested by significance testing and different patterns are found among never, former and current smokers but not referred correctly in the text.
29. We have added to the results section (pages 12 and 15) more explicit mention of statistical significance when results in different cohorts, or in different smoking subgroups, are being compared.

The manuscript had been easier to follow if there were more specific aims, and statistics and results were reported in the same order.
30. The introduction (page 5-6) has been extended to clarify the aims (see also point 5 above). We are not sure what is meant by the second part of this sentence, but we hope that point 29 above addresses the concern about aligning statistics and results.

Comparisons between different qualities of spirometry and between the measuring sites are of interest, but shouldn't rather the comparisons with other measures like BMI and BP have been given priority. The unexpected U-shape for BMI and BP compared to linear relation between quartiles of z-scores should be discussed more thoroughly.
31. The focus of this paper is on ventilatory function and the comparisons with BMI and SBP were included mainly for context, and partly because they had been compared with FEV1 as predictors in one of the two previous studies (reference 4). A U-shaped relationship of all-cause mortality with BMI is widely recognised, but we agree that the U-shaped relationship for SBP is unexpected, particularly as it is also seen for circulatory mortality in UK Biobank. However, we feel strongly that this is topic of another paper and it would be a substantial digression to explore it fully here.

Comments:
Abstract line 22-24: GLI-2012 z-scores are standardized measures not standard deviations. Among causes of death respiratory diseases should also be included.
32. We acknowledge the terminological point and have changed "per SD decrement" to "per unit decrement" throughout, when referring to z-scores. We have included results for respiratory deaths in the abstract (p3) as requested also by referee 2 (point 19 above).

Introduction: The authors write that the HSE and SHS are more representative of the general population than UK Biobank. What was the participation rate in the age group 40-69 years in HSE/SHS? In what way did the participants differ between HSE/SHS and UKB? Have any nonparticipation studies been performed?
33. No information is available on the participation rate specifically for the 40-69-year-olds in the national health surveys. Response rates for all ages are published in the survey reports and summarised in the following two papers, which confirm that the participation rates in HSE and SHS increase with age and are substantially higher than in UK Biobank. We are not aware of any studies which have compared participants with non-participants in either HSE-SHS or UK Biobank.
Mindell J, Biddulph JP, Hirani V et al. Cohort Profile: the Health Survey for England. Int J Epidemiol 2012;41:1585-1593.
Mindell JS, Giampaoli S, Goesswald A et al. Sample selection, recruitment and participation rates in health examination surveys in Europe – experience from seven national surveys. BMC Medical Research Methodology 2015;15:78.
A major difference between the participants in HSE-SHS and UK Biobank is the distribution of smoking habits, now presented in the supplement (see point 34 below). The proportion of lifelong non-smokers is much greater in Biobank, but this is an advantage in the context of our study, in which the main manuscript focuses on never-smokers.

The statistics should be described in more detail: There should be a table reporting descriptive data of the study samples being compared; All, UKB, HSE-SHS, those with non-respiratory mortality, circulatory mortality etc.
34. This information has been included in the supplement (new e-Table 1) because providing descriptive data for each of the three smoking subgroups fits better with the format of the supplementary tables. The relevant information for never smokers was already included in the text in the original submission. This supplementary table is mentioned at p10 in the text.

The GLI z-scores are standardized measures of lung function adjusted for age, sex and height. If there is lack of fit for GLI-2012 in the studied population, the z-scores could vary by these confounders. Is this the rationale for adjusting for age, sex and height in the proportional hazard models?
35. This is indeed our rationale.

In many tables HR are reported and compared; in the statistics it should be stated whether

assumptions were violated or met. The tables report HR from many models; how has the authors tested HR from different models?
36. The validity of the proportional hazards assumption for each of the smoking subgroups is addressed in detail in annex 2 (see also point 15 above).

The legend of table 1 should include by GLI Z scores for FEV1, FVC and FEV1/FVC. According to the footnotes the HRs are expressed per SD decrement in z-score; this is not correct.
37. We accept this technical point and have changed “per SD decrement” to “per unit decrement” throughout, including the supplementary table footnotes. The legend for Table 1 (page 11) has been extended as suggested.

Page 11 line 11; what is meant by the most direct comparison? I cannot see that the statement: Although the associations of spirometric indices with these earlier deaths were stronger in HSE-SHS than in UKB…. is correct according to data given in e-table 1. Further, how was significance of differences tested? I cannot see that the statement in line 27-31 is in accordance with data in e-table 2.
38. The rationale for comparing deaths within 5 years is now explained on page 10. The formal statistical tests of time-dependence of the hazard ratios for all-cause mortality in never, former and current smokers are now presented on page 12. The statistical significance of the comparison between cohorts is also now included on page 12.

Page 14 line 20-25. The text regarding male never smokers cannot be correct. And slightly stronger association than??
39. This sentence has been reworded for clarity (page 15).

Page 14 line 35-42. Is it correct that the pattern of results was consistent between the age categories in all smoking subgroups?
40. This section has been elaborated to provide more detail on the pattern of results in each smoking subgroup, within each cohort (page 15). There is a significant age interaction for both spirometric indices among ex-smokers in the national health surveys. Otherwise, the tests for effect modification by age are non-significant.

VERSION 2 – REVIEW

REVIEWER	David Batty UCL
REVIEW RETURNED	12-Apr-2017

GENERAL COMMENTS	The authors have responded comprehensively and thoughtfully to my observations.

REVIEWER	Joachim Heinrich Helmholtz Zentrum Munich, Institute of Epidemiology I, Germany
REVIEW RETURNED	30-Apr-2017

GENERAL COMMENTS	The manuscript substantially improved after revision. The authors responded well to my comments.