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)	The burden and impact of chronic cough in UK Primary Care: a dataset analysis
AUTHORS	Hull, James; Langerman, Haya; Ul-Haq, Zia; Kamalati, Tahereh; Lucas, Amanda; Levy, Mark

VERSION 1 – REVIEW

REVIEWER	Tiejun Zhang Fudan University School of Public Health, Epidemiology
REVIEW RETURNED	09-Sep-2021

GENERAL COMMENTS	In this manuscript, the authors have detailed about the basic characteristics, comorbidities and service cost of chronic cough based on the primary and secondary care data set. They found that 1/3 individuals had CC in the absence of associated comorbidities, and the outpatient costs increased in the year after the CC index date for all comorbidities. This analysis may provide some useful information for the management of CC in this future. I have some questions for the authors. 1. The dataset last from 2015 to 2019, has this definition of CC been changing? Thus affect the report of the CC case. 2. Some comorbidities are hard to detect and difficult to tell which one in the first, the comorbidity or the CC. 3. There is a significant discrepancy in the CC prevalence between the present study and previous reports, could the author elaborate on this issue. Is there any other reasons besides they have mentioned in the text? 4. Information in the table 1 and Figures may be redundant.

REVIEWER	Kerry-Ann O'Grady Queensland University of Technology, Institute of Health & Biomedical Innovation
REVIEW RETURNED	21-Sep-2021

GENERAL COMMENTS	This study describes a large administrative data-derived analysis of the prevalence of chronic cough in a region of the UK. CC is an important cause of morbidity, health care costs and decreased quality of life globally however, as the authors point out, until recently, standardising the definition of chronic cough in prevalence studies has been suboptimal. Overall the paper, albeit descriptive, is a useful contribution and my comments are relatively minor and for clarification purposes rather than any major changes, apologies if I have misinterpreted something. 1. "cough-related" coding. Kindly clarify if the record had to have at

least one of the "Cough" codes presented in supplementary document 1. I am not aware of the use of this dataset but is it standard for doctors to code a symptom in addition to a diagnosis such as lower respiratory tract infection. Further, did a "cough" code have to present at both time 1 and time 2? If possible (albeit understandably having limitations), some indication of how frequently a cough code was present at either, but not both, timepoints, particularly in those with relevant comorbidities would be of interest.

2. "A prescription for cough related remedy" as one of your criteria. Was this prescription required at both T1 and T2 or just one? If just at T2 (and a cough related code was not included), it could have been for the severity of an acute cough, not the duration of the cough

3. I am not sure why, given the dataset you had, you did not compare your outcomes between subjects with CC and comorbidities and those with the same cormorbidities but no CC (or only 1 coded timepoint with cough). This may have given further insight into the additional contribution of cough beyond a "before and after analysis". It would also have given an estimate of the prevalence of CC with comorbidities in addition to the estimate of the prevalence of comorbidities in those with CC

4. I am not sure of the justification for excluding in-patient episodes and costs. Is this because it would unclear if the admission was due to the cough or the underlying illness? It would be worthwhile mentioning that either in the methods or in the discussion.

5. The completeness of data around smoking status is very low, surprisingly so for this particular cohort. Some comment around that in the discussion would be useful. Does the dataset collect anything on referral to smoking cessation services? If not OK, but if it does it would be interesting to see how many were referred as part of their management given smoking cessation is an important part of the management plan.

6. The reference to the Cullinan study. It would be helpful to have the prevalence reported in that study included here..while I understand it is on page 1, I had to check what the Cullinan paper was...either that or put the name on page 1 in the text.

7. Reference to the CHEST cough management guidelines/algorithms for adults that were update in 2017 is needed. Your reference #3 refers to the 2006 version.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Tiejun Zhang, Fudan University School of Public Health Comments to the Author:
In this manuscript, the authors have detailed about the basic characteristics, comorbidities and service cost of chronic cough based on the primary and secondary care data set. They found that 1/3 individuals had CC in the absence of associated comorbidities, and the outpatient costs increased in the year after the CC index date for all comorbidities. This analysis may provide some useful information for the management of CC in this future. I have some questions for the authors.

1. The dataset last from 2015 to 2019, has this definition of CC been changing? Thus affect the report of the CC case.

The primary care computerised data collection dates back to the 1980's, but was only linked to the Discover dataset from 2015 onwards, hence the choice to utilise 2015-2019 for the study period. The definition we used for Chronic Cough is defined in the first sentence of the introduction (lines 71-72). In this study, patients with Chronic Cough were defined as those with two or more recorded consultations coded as 'cough-related' persisting for at least 8 weeks, and this definition was unchanged throughout the study (clarified in lines 142-144 and we have amended the text in the Study Design section (lines 157-158).

2. Some comorbidities are hard to detect and difficult to tell which one in the first, the comorbidity or the CC.
The primary care records which are contemporaneous and in chronological order provide the ability to determine which came first, however, we have not included these details in our analysis. We have added a note to this effect in the study limitations (lines 359-362).

3. There is a significant discrepancy in the CC prevalence between the present study and previous reports, could the author elaborate on this issue. Is there any other reasons besides they have mentioned in the text?
There are a number of potential reasons for variations in prevalence between studies. These include the definition of chronic cough used to identify individuals and the age of individuals included. We have now addressed this important point in the text in line 297 of the Discussion.

4. Information in the table 1 and Figures may be redundant.
Table 1 presents the demographic and patient characteristics. While some of these data are illustrated in Figures 2 and 3 the authors believe these help the reader understand the data. We would, however, be happy to be guided by an editorial decision on whether the table and figures duplicate data and thus should be adjusted accordingly.

Reviewer: 2
Dr. Kerry-Ann O'Grady, Queensland University of Technology Comments to the Author:
This study describes a large administrative data-derived analysis of the prevalence of chronic cough in a region of the UK. CC is an important cause of morbidity, health care costs and decreased quality of life globally however, as the authors point out, until recently, standardising the definition of chronic cough in prevalence studies has been suboptimal.

Overall the paper, albeit descriptive, is a useful contribution and my comments are relatively minor and for clarification purposes rather than any major changes, apologies if I have misinterpreted something.

1. "cough-related" coding. Kindly clarify if the record had to have at least one of the "Cough" codes presented in supplementary document 1. I am not aware of the use of this dataset but is it standard for doctors to code a symptom in addition to a diagnosis such as lower respiratory tract infection. Further, did a "cough" code have to present at both time 1 and time 2? If possible (albeit understandably having limitations), some indication of how frequently a cough code was present at either, but not both, timepoints, particularly in those with relevant comorbidities would be of interest.

Yes at least 1 of the cough codes had to be present at both the first and second visits, ie to persist for at least 8 weeks to satisfy the Chronic Cough definition used in this study. This has been added to the Study Design (lines 157-158).

Since 1983/4, General Practitioners in the UK have entered consultation data into a computerised record. There is some variance in the quality of data entry, with differences from one general practice to another, some entering details of both symptoms and diagnosis. The limitations of utilising this type of data for research are described in the Discussion (lines 350-359).

A Cough code had to be present at time 1 and 2 in order to satisfy the definition of Chronic Cough. We did not analyse the frequency of cough codes between these time points and have added a comment as a limitation of the study in the Discussion (lines 362-363).

3

2. "A prescription for cough related remedy" as one of your criteria. Was this prescription required at both T1 and T2 or just one? If just at T2 (and a cough related code was not included), it could have been for the severity of an acute cough, not the duration of the cough
The prescriptions had to be present at T1 and also T2 to satisfy meeting the definition of CC ie with a duration of at least 8 weeks. We have added a note to this effect in the methodology under Study Design (lines 157-158). We have also amended the legend for Figure 1 (lines 536-537), which we prefer to retain in the document.

3. I am not sure why, given the dataset you had, you did not compare your outcomes between subjects with CC and comorbidities and those with the same cormorbidities but no CC (or only 1 coded timepoint with cough). This may have given further insight into the additional contribution of cough beyond a "before and after analysis". It would also have given an estimate of the prevalence of CC with comorbidities in addition to the estimate of the prevalence of comorbidities in those with CC
We have compared the costs between those with and without comorbidities before and after the CC index date and summarised these in Figure 4. The text has been amended to clarify this (lines 238-239). The legend for Figure 4 has also been amended (lines 544-545).

4. I am not sure of the justification for excluding in-patient episodes and costs. Is this because it would unclear if the admission was due to the cough or the underlying illness? It would be worthwhile mentioning that either in the methods or in the discussion.
In-patient episodes and costs were excluded because the Discover dataset only has access to the ICD-10 coded discharge diagnosis which wouldn't clarify whether chronic cough was the reason for the admission. Therefore, it would be unclear if the admission was due to the chronic cough or the underlying condition. To address this point for readers we have added additional detail in the Study Design section (lines 169-171).

5. The completeness of data around smoking status is very low, surprisingly so for this particular cohort. Some comment around that in the discussion would be useful. Does the dataset collect anything on referral to smoking cessation services? If not OK, but if it does it would be interesting to see how many were referred as part of their management given smoking cessation is an important part of the management plan.
The dataset does cover Referral to Smoking Cessation Services. Although it would be potentially interesting to know how many patients were referred to Smoking Cessation Services, this would not indicate if the patients attended or acted on the advice. For this reason, we chose to actively incorporate patients who actually had the codes for Smoking Cessation (EX-smokers).

In lines 298-299 we state that one in five (20%) of our Chronic cough cohort were either current or ex-smokers, which is higher than the 14.1% national prevalence for smoking in the UK (see lines 298-300) Smoking status was retrieved for 44.2% of the CC cohort. It is unclear why it was not recorded in more individuals. (Please see lines 351-359 where we have commented on the limitations of utilizing retrospective computerized data in general practice records).

6. The reference to the Cullinan study. It would be helpful to have the prevalence reported in that study included here..while I understand it is on page 1, I had to check what the Cullinan paper was...either that or put the name on page 1 in the text.
We have now added the author name where reference 14 is mentioned on page 4 (line 95). The prevalence rates in the Cullinan study have been added to the Discussion (lines 296-297).

7. Reference to the CHEST cough management guidelines/algorithms for adults that were update in 2017 is needed. Your reference #3 refers to the 2006 version.
Thank you for pointing this out. The 2018 paper confirmed that the assumptions from 2003 were still valid although they updated the algorithms. We have added clarifying text on p.3 (lines 77-78) and added the reference (lines 438-440).

VERSION 2 – REVIEW

REVIEWER	Tiejun Zhang
	Fudan University School of Public Health, Epidemiology

REVIEW RETURNED	24-Nov-2021

GENERAL COMMENTS	I have no further comment.

REVIEWER	Kerry-Ann O'Grady Queensland University of Technology, Institute of Health & Biomedical Innovation
REVIEW RETURNED	25-Nov-2021

GENERAL COMMENTS	Thank you for addressing my comments clearly and succinctly. It is a nice paper that is a worthwhile addition to the literature