Author’s response to reviews

Title: Community Assessment of COPD Health Care study: a clinical audit on primary care performance variability in COPD care

Authors:
Jose Luis Lopez-Campos (lcampos@separ.es)
Maria Abad-Arranz (marieta.84@hotmail.com)
Ana Moran-Rodríguez (anamoran@comcadiz.com)
Enrique Mascarós Balaguer (enmasba@gmail.com)
Carmen Quintana Velasco (cquintanav66@gmail.com)
Laura Abad Polo (laura.abad.polo@gmail.com)
Sara Núñez Palomo (snpalomo@sociedadgrap.com)
Jaime Gonzálves Rey (gonzalvez.jaime@gmail.com)
Ana Maria Fernández Vargas (anamorronga@gmail.com)
Antonio Hidalgo Requena (ahidalgor1@yahoo.es)
Jose Manuel Helguera Quevedo (josehelg@hotmail.com)
Marina García Pardo (marinagpardo@hotmail.com)

Version: 2 Date: 21 Mar 2018

Author’s response to reviews:

Terttu Harju (Reviewer 1):

General comments:

To audit current dg and treatment of common preventable and treatable diseases such as COPD in real life is an important area of research. Provided local and global guidelines do not work without careful implementing and audit with benchmarking according to given guidelines helps to focus further education and resources.

This is a retrospective study auditing the diagnosis and care of stable COPD in 63 randomly selected primary care centers in Spain in 6 regions, with 4307 patient cases, during a time period
of two years (1.1.2015-31.12.2016). The standards of diagnosis and care were set by two well-established guidelines, the Spanish COPD guidelines and the global GOLD 2017 guideline.

The results reflect the real world in primary care setting: the diagnosis of COPD is rarely based on the diagnostic criteria and standards described in guidelines. This is an important area for education. Even though the diagnosis of COPD is not based on diagnostic testing, COPD medication is widely used, especially inhaled corticosteroids and this makes me wonder if there is a significant misdiagnosis and overtreatment / mistreatment of chronic bronchitis vs COPD. This should be discussed in detail and recommendations for clinicians given, based on current audit. Also educational goals aiming to improve diagnosis and treatment of COPD could be set. Maybe also the GOLD guidelines for COPD dg (exposure, symptoms, post-bd spirometry FEV1/FVC <70%) as well as treatment flow chart could be cited in a information box, for those readers not so familiar with COPD.

RESPONSE: We thank the reviewer for the positive comment. As you also mentioned, these are important points regarding the diagnosis of COPD. In the light of the audit results, we have included a set of recommendations in the discussion and the diagnostic guidelines in a figure (figure 1), as per the suggestion.

Minor comments:

Page 3, row 15: the term randomized is misleading; randomly selected would be a more precise term.

RESPONSE: We have changed this term throughout the manuscript, as per your suggestion.

Page 7, row 32: case definition should be stated: how was a case "COPD" defined

RESPONSE: We thank the reviewer for this relevant comment. The procedure by which COPD was defined is explained on page 10: “the local investigator consecutively reviewed the medical records for confirmed cases of COPD and recorded the audit information for 80 cases per PCC”. Accordingly, we selected cases that had a diagnosis of COPD in the medical record. This is a
crucial point as one of the aims of a clinical audit is to evaluate the correctness of the diagnosis itself. To do so, we selected those cases with a diagnosis in the medical record and then evaluated the diagnostic criteria. We have added a comment to this point for clarification.

Page 7, row 56: classified (not casssified)

RESPONSE: Thank you for pointing out the typographical error. We have reviewed the whole document to avoid any spelling mistakes.

Page 8, rows 34-43: asthma symptoms (not asthma-like symptoms) and either to 'described in GINA guidelines (ref) ' or wheezing ..... (ref Gina)

RESPONSE: We have corrected the term and included the reference for GINA.

Page row 47: inhaler satisfaction??? Does this refer to correct inhaler technique or to patient preference / satisfaction to inhaler device? Clarify please.

RESPONSE: The reviewer is right by pointing this out. This is not a very strong variable. Since this was a clinical audit, the clinicians involved in the steering committee felt that the general satisfaction of the patient with the inhaler should also be evaluated. This was further discussed during various meetings and during the development of the item selection. This variable does not specifically reflect the correct technique for inhaler use or patient preference, but if some information indicating the general satisfaction of the patient with the inhaler was recorded. I understand that this may include different aspects of the inhaler, e.g. comfortability, ease of use, patient preference, number of inhalations per day, and so on. We have included an explanation in the methods section to clarify this point.

Page 9 row 7: the diagnosis of COPD should be on post-bd spirometry, not pre-bd. The number of cases with post-bd spirometry FEV1/FVC <70% should be stated (xxx cases spirometry was
available; xxx cases also post-bd spirometry and xx% of post-bd FEV1/FVC was less than 70% confirming the COPD-diagnosis)

RESPONSE: The reviewer poses a relevant question here. We fully agree with the reviewer that a correct diagnosis of COPD should only be based on post-bronchodilator spirometry. The analysis plan, however, considered a probable COPD patient with symptoms and a pre-bronchodilator spirometry with obstructive pattern. We know this is not strictly correct, but since this is a clinical audit, there were cases with a diagnosis of COPD based on post-bronchodilator spirometry and a more recent pre-bronchodilator spirometry. For example, there were cases with a previous post-bronchodilator spirometry with obstruction, but the most recent test recorded was only a pre-bronchodilator spirometry. Or cases with a pre-bronchodilator spirometry but no asthma symptoms and a severe obstruction. Additionally, only 67 cases had a full post-bronchodilator spirometry with obstruction in the audited visit, which will under-estimate the actual diagnosis rate. To clarify the point and to provide the complete information, we have mentioned in table 2 what type of obstruction we were referring to and have included the requested information in the results section. Additionally, we have added the number of cases with correct diagnosis to the text, considering only those cases with post-bronchodilator spirometry, which is considerably lower.

Page 12 row 14: some centers improved as the study progressed - this should be discussed. So this study was not entirely retrospective? Was it partly retrospective and partly 'on time’? What was the time span between starting the audit and the time period under evaluation?

RESPONSE: We thank the reviewer for noticing this and we understand the confusion. What we meant was that some centers did better than others. There was no prospective follow-up phase. We have modified the text to make it clearer.

Page 13 row 9: Only 10% of cases could be classified into GOLD classes A-D. Given treatment was analysed in this subgroup. The possible bias should be discussed. Was this subgroup similar
to total material? Females, age, lung function, fewer comorbidities etc. compared those without enough data to be classified.

RESPONSE: This is an interesting question. We had not done this analysis initially. We have now explored this, and indeed there are some significant differences between those cases with and without a complete GOLD categorization information. We are not sure, however, whether these differences are relevant in terms of data interpretation, since they are clearly influenced by the considerable sample size, where small differences may be statistically significant. For example, cases with a complete GOLD 2017 information presented some differences with the rest of the cohort, e.g. they were more frequently male (79.1% vs 73.3%; p=0.009), younger (69.8 vs. 71.3 years; p = 0.017), more frequently current smokers (28.9% vs. 18.0%; p < 0.001), and had a higher Charlson index (2.3 vs 2.2; p = 0.043), but these little differences are probably not clinically relevant. We have included a comment and highlighted that only 10% of cases could be categorized according to GOLD 2017. This has been added to the Discussion.

Discussion: the poor implication of current guidelines should be discussed. The areas of greatest concerns should be emphasized: only correct diagnosis makes it possible to select proper pharmacological treatment for COPD. COPD cannot be diagnosed without proving irreversible obstruction (according to GOLD guidelines). Cigarette smoke exposure and especially duration of smoking / years is an important risk factor. Smoking cessation, physical activity, nutrition, vaccinations as well as recognition and treatment of comorbidities are even more important than bronchodilators. Golden standard for treatment of chronic bronchitis without obstruction is not there yet. The use of ics (42.2% in this material!) in COPD patients should be confined to those who do get the benefit but no harm.

RESPONSE: We fully agree with the reviewer in all these points. We have added a completely new paragraph in the discussion on page 22 summarizing all these valuable points as recommended.
Table 1. Average (patient level)???. Number of patients (%) could be used instead. Smoking status: should be presented more accurately: active smokers / ex-smokers / life-long nonsmokers / smoking status unknown. OR smoking status known xxxx and xx% current smokers. Pack-years 365 ????. FEV1/FVC should be added to the table. The percentage of data post-bd vs pre-bd should be stated (for FEV1, FVC and FEV%).

RESPONSE: We thank the reviewer for this thorough evaluation of table 1. We want to draw your attention to a few points in this regard. The statistic that evaluates results in health based on hierarchical data has become remarkably advanced in recent years. In databases with a leveled structure, like ours, it is necessary to identify the stratum on which the analysis is centered. For this reason, it is necessary to reflect whether the data analyzed is at the patient level or higher. As a result, we have marked this level in the tables. The reviewer is correct when categorizing smoking status that way. We have included the complete information in table 1. Thanks for noticing the mistake in the pack-years, we have now corrected it. We have added FEV1/FVC to the table, as per your request.

Table 2. exposure should be defined (for example in the footnote). Asthma symptoms present (not asthma-like symptoms; my preference). The term 'recorded': does it refer to the presence of for example habit of exercising or pneumococcal vaccine taken or that the information is available, either +/-? Solicited???. Could a more precise term be used? Please clarify. Is this 'Complementary tests' section needed or could it be deleted?

RESPONSE: We have changed the asthma term, as per suggestion. As this is a clinical audit, the term recorded refers to whether this information appeared in the patient's medical record or not. For example, we asked if information about exercise was found in the medical record, not whether the patient actually exercised. There is a complete section to explain this in the discussion. Maybe a better term could be registered. We have made this change throughout the document. Equally, because it is an audit, it is important to record the clinical performance, which includes the use of complementary tests.
Figure 2 and 3. Add to figure legend or a footnote: In xxx cases out of total xxx GOLD classification could be performed, based on exacerbation frequency and symptoms (cat or mrc).

RESPONSE: We have added this sentence as a footnote for figure 4, as per the suggestion.

Stefan Markun, M.D. (Reviewer 2):

The authors are to be commended for their relevant work because most of the patients affected by COPD are treated in general practice and consecutively depend on the quality of care delivered in this healthcare setting. The manuscript describes an audit performed to assess the degree of adherence to guidelines for stable COPD in Spanish general practice. The manuscripts merits consideration because of the relevant topic but in my view, the following points should be addressed by the authors:

RESPONSE: We thank the reviewer for the positive comment. We have provided our responses to each one of the queries below.

ABSTRACT

- "To the best of our knowledge, there has been no previous audit of primary health care interventions for chronic obstructive pulmonary disease (COPD) patients.": Several studies have assessed performance of primary healthcare for COPD (see: Jochmann A et al.: General practitioner's adherence to the COPD GOLD guidelines: baseline data of the Swiss COPD Cohort Study. Swiss Med Wkly 2010; Steurer-Stey C, et al.: Management of chronic obstructive pulmonary disease in Swiss primary care: room for improvement. Qual Prim Care 2012; 20: 365-373.; Kaufmann C et al.: Performance Measures in the Management of Chronic Obstructive Pulmonary Disease in Primary Care - A Retrospective Analysis. PRAXIS 2015; 104: 897-907.; Belletti D et al.: Results of the CAPPS: COPD - Assessment of Practice in Primary Care Study. Curr. Med. Res. Opin. 2013; 29: 957-966.). Even if they are not labelled as "audits" these studies reflect performance of primary care for COPD. In addition a simple google search inputting "audit COPD primary care" identified two more studies on the first search page, one being a very
large project from Wales UK (https://www.rcplondon.ac.uk/file/8467/download?token=u8x9UeTT). The opening statement thus seems inaccurate and should be amended.

RESPONSE: The reviewer is correct; there are several studies that have assessed performance of primary healthcare for COPD. As the reviewer mentioned, although these are not formal clinical audits, they provide valuable information on clinical performance in this setting, but use different methodologies. Therefore, we still believe that ours is the first ad-hoc formal clinical audit evaluating COPD care in primary care settings. In order to avoid making absolute statements and following the reviewer recommendation, we have changed this sentence in the abstract, and included some of the most recent studies on this topic in the introduction.

- "The Community Assessment of COPD Health Care (COACH) study was an observational, multicenter, nationwide, randomized, non-interventional, clinical audit of primary care in Spain." Please state also whether the audit was performed prospectively (with an associated comment in Limitation section concerning awareness of physicians risk for over-estimation of true performance [Hawthorne-Effect]) or retrospectively (with associated comment in Limitation section concerning unawareness of physicians risk of under-estimation of true performance because of lacking documentation about interventions performed in reality i.e. smoking cessation advice performed but not documented [information bias]).

RESPONSE: The study was retrospective. We have added this in the abstract and methods section. As we state in the methods section “The doctors in charge of the patients were unaware of the audit”, and there is a comment on this potential bias, which the reviewer correctly pointed out, on page 15.

- "Diagnosis based on previous exposure plus...": By exposure you mean exposure towards risk-factors? If yes I don't understand because risk-factors are not a requirement for diagnosing COPD (yet they are a case-finding strategy to decrease the number of patients with unidentified COPD and may be used for screening for COPD [but not for diagnosis, especially because
COPD rarely also occurs without exposure to risk factors, i.e. alpha-1-antitrypsin deficiency). In my understanding, diagnosis of COPD should rest on spirometry including reversibility testing.

RESPONSE: We fully agree with the reviewer that COPD should rest on spirometry including reversibility testing. As the reviewer mentioned we mean exposure to a risk factor, since the current GOLD document includes “exposure to noxious particles or gases” as part of the definition of the disease. Additionally, in the diagnosis of COPD, the GOLD document further specifies that “COPD should be considered in any patient who has dyspnea, chronic cough or sputum production, and/or a history of exposure to risk factors for the disease. Spirometry is required to make the diagnosis in this clinical context”. Therefore, we believe that exposure to risk factors is a requirement for the definition and diagnosis of the disease, although the final diagnosis will then be based on spirometry.

- "Regarding treatment, 33.6% received no maintenance inhaled therapies.": This result is pure description without questioning the appropriateness of the intervention (what I would expect from an audit). Do the guidelines you operationalized not provide recommendations on who to treat with inhaled therapies? Since you state to have reports about exacerbations within previous year it should be feasible to check whether at least those with exacerbations received inhalative treatments. This would enable you to distinguish the under-treated population where improvements are possible. This would make the results meaningful in the context of the manuscript's aim to be an audit.

RESPONSE: We thank the reviewer for this comment and letting us clarify. We agree that giving the isolated overall data does not garner much interest. It is possible that some of these patients did not receive treatment because they did not need it. To evaluate this in relation to the clinical presentation we would like to refer you to the information presented in figure 4, in which you can see these same numbers divided by patient types based on the GOLD document. Specifically, cases with no maintenance inhaled therapies were: GOLD A 31.3%, GOLD B 8.1%, GOLD C 9.1%, GOLD D 7.0%. We have updated the abstract and included this information in the results section.
BACKGROUND

- page 5, line 15: check writing "...are currently have shown..."

RESPONSE: Thanks for noticing this, we have now corrected it.

- page 5, line 56-59 "However, to the best of our knowledge, there has been no previous audit of primary health care interventions for COPD patients.": Please amend according to comment on same statement in Abstract

RESPONSE: We have made these corrections. There are quite a number of valuable studies evaluating clinical performance but with different methodologies. The current Welch Audit mentioned (which is a real clinical audit) has not been published in a scientific journal so far. Hence, strictly speaking, it is true that this would be the first publication of a COPD clinical audit in primary care in a scientific journal. We understand, however, that there are previous publications. Hence, we have modified the text accordingly, so as not to be so imperative.

METHODS

- Spelling error detectable by Microsoft Word (page 7 line 56 "calssifed")

RESPONSE: Thank you for pointing this out. It has been corrected.

- page 8 line 27: were the number of exacerbation the audited measure or was the processes of ASSESSING this number audited (logically I can think only of the second [because assessing the number is recommended by guidelines to enable steering of care / selecting appropriate interventions] but it reads as the first would be the intention of the audit which would reflect just the disease severity in the population but not the quality of care)

RESPONSE: The reviewer is correct. Assessing the number is recommended by guidelines to select appropriate interventions. We have added this in the text to make it clear.
- page 8 line 53: as commented before, the previous inhalation exposure is not a diagnostic requirement for COPD. This seems to belong to a different study question, namely whether testing for COPD is appropriate. It would be interesting whether the GPs actually consider COPD in the population with inhalation exposure and the proportion of this population actually identified and forwarded to spirometry. On the other hand, what proportion of patients receives spirometry without risk factors. Together these proportions would estimate the failure to test for COPD (under-use) and unnecessary tests for COPD (over-use) which would be the meaningful audit outcomes. The data, however, just grasps positively identified COPD cases leaving unknown the proportions were diagnostic testing would have been needed but not performed. Therefore, if my understanding of COPD diagnosis is accurate then I would detach the inhalation exposure from the accuracy of diagnosis definition. Otherwise I would welcome an update on my understanding of COPD diagnosis by the authors.

RESPONSE: Thanks for pointing this out. We have commented on this aspect in comment number 4. According to the current GOLD document, “exposure to noxious particles or gases” is part of the definition of the disease. Additionally, in the diagnosis of COPD, the GOLD document refers that “COPD should be considered in any patient who has dyspnea, chronic cough or sputum production, and/or a history of exposure to risk factors for the disease. Spirometry is required to make the diagnosis in this clinical context”. Therefore, we believe that exposure to risk factors is a requirement for the definition and diagnosis of the disease, that will then be finalized based on spirometry.

- page 9 line 7. Accepting also spirometries without bronchodilatator is pragmatic and may have moderate false-positive inclusion because of asthma. however, it hardly reflects the guidelines saying which you declare to use as standards of care. Why do you make this trade-off, would it not be very important to report the appropriateness of diagnostic procedures this in your audit?

RESPONSE: The reviewer poses a relevant question here. We fully agree with the reviewer that a correct diagnosis of COPD should only be based on post-bronchodilator spirometry. The analysis plan, however, considered a probable COPD patient with symptoms and a pre-bronchodilator spirometry with obstruction. We know that this is not strictly correct, but, since
this is a clinical audit, there were cases with a diagnosis of COPD based on post-bronchodilator spirometry and a more recent pre-bronchodilator spirometry. For example, there were cases with a previous post-bronchodilator spirometry with obstruction, but the most recent test recorded was only a pre-bronchodilator spirometry. Or there were cases with a pre-bronchodilator spirometry but no asthma symptoms and a severe obstruction. Additionally, only 67 cases had a full post-bronchodilator obstructive spirometry in the audited visit, which will also under-estimate the actual diagnosis rate. To clarify the point and provide the complete information, we have mentioned in table 2 what type of obstruction we were referring to and we have included the requested information in the results section. Additionally, we have added the number of cases with correct diagnosis, considering only those cases with post-bronchodilator spirometry, which is considerably lower.

- I cannot find how cases were identified. Electronical searches or manual searches of consecutive patients? Also: who identified the cases? The doctors in charge or the person conducting the audit?

RESPONSE: As stated in the methodology, the local investigator consecutively reviewed the medical records to find those cases with a confirmed diagnosis of COPD in the medical record and record the audit information.

- Do "cases" translate into "patients" implying that every individual contributes to the dataset only once or is there a risk of single patients being counted/audited multiple times? If yes this should be mentioned in the Limitations section. If no, what precautions were in place, given that the audit form used did not contain personal data as described on page 10 line 43. In summary, the methods lack a description how audited cases were identified, usually a flow-chart is used to further illustrate this process for readers.

RESPONSE: We thank the reviewer for letting us clarify. Cases translate into patients. The local investigator ensured that duplicate reports for the same patient were not made. As stated in the ethical section, the relationship between the audit code and the clinical history number was kept
locally and was the local investigator’s responsibility. We have added a short comment here to clarify.

- The Methods section should be re-written and I strongly recommend using a reporting guideline in order to address all necessary points of methodology.

RESPONSE: We would like to ask the reviewer to be more specific. We have answered all the queries regarding methodology and updated the content accordingly. We have revised the equator initiative in search of reporting guidelines and there are none for clinical audits in COPD patients. Instead, we have followed the same structure as in previous clinical audits regarding COPD patients. We believe that the methodology is complete with all the information requested by the reviewer included.

RESULTS

- No results from the process of identifying audited cases are visible. This would be very important to understand representativeness of the population starting from the participation rates of physicians. Please include this results from the process of accruing your study population on all levels (ideally in the flow-chart).

RESPONSE: The audit aimed at exploring diagnostic accuracy and therapeutic approach for those cases labelled as COPD. Therefore, there were no strict inclusion or exclusion criteria as in a traditional observational study or a randomized clinical trial that could be represented in a flowchart. The entry criterion was a current diagnosis of COPD in the medical record. The appropriateness of these criteria was part of the audit. As stated in the methodology, the local investigator consecutively reviewed the medical records to find those cases with a confirmed diagnosis of COPD in the medical record and record the audit information.

- Page 12 line 2: Since you report on variability concerning various patient characteristics and find significant differences i expect in the Discussion section a comment on the implications and
relevance of these numerous differences especially with regard to COPD care. If you see no important implications even if you found so many differences, then I would wonder why you have measured and tested for all these differences in the first place.

RESPONSE: Indeed, variability is key in clinical practice and one of the key goals of clinical audits, that is why we are reporting it. We would like to thank the reviewer for letting us clarify. Previous studies have demonstrated significant variability in the processes of care and outcomes of COPD care in different settings. It has been shown that this variability in clinical practice is not exclusively influenced by disease severity, clinical presentation, or the resources available. In fact, a so-called cluster effect is described. The cluster effect indicates that patients with similar characteristics may experience different processes of care and outcomes, depending on the center because they are subject to distinct common contextual influences, beyond resources or standards of care, like the characteristics of the catchment area, such as socioeconomic status and utilization of health services or COPD specific criteria in a determined clinical situation. In our audit, we found a wide range of variability in clinical practice that should be explored in the future. The idea behind quality of care is that all patients with the same condition should be managed similarly. However, this is very complex in clinical practice since clinical practice is influenced by a number of additional factors, including the variation in the clinical presentation of chronic diseases; the perception of the patient, which is influenced by many personal aspects; and the varied responses to therapies. Therefore, describing the variability is essential to really understand its potential implications, allowing us to seek for specific approaches in the future.

We have added a comment about this extremely interesting debate in the discussion, as per the suggestion.

- Page 12 line 2: Apart from this comment belonging to the Discussion section: apart from gender, age and BMI, FEV1 is one of the few anthropometric measures among the patient characteristics you compare. There is an important risk for diagnostic bias in antropometric measures but even more in the comorbidities you report. I believe it to be very difficult to distinguish actual morbidity of the population from mere performance of healthcare in establishing the diagnoses you report. Therefore I recommend not to over-interpret this
variability in patient characteristics which are also hardly an important outcome of your audit on COPD healthcare.

RESPONSE: We are not sure how to interpret the reviewer’s point here. The reviewer is mixing gender and age, with an anthropometric measure (BMI) with a lung function parameter (FEV1) and with comorbidities. As stated in the methodology section, these variables (except for FEV1) are recorded for descriptive purposes of the patients included in the audit.

- Page 12 contains many comments on results ("Interestingly..."). These belong to the Discussion section.

RESPONSE: The reviewer is correct. We had allowed ourselves this small lapse because we felt that the reader could appreciate some interpretation of these interesting results. We have now moved the comments to the discussion section, as per the suggestion.

- Page 13 line 15: To prescribe ICS in GOLD A and B I understand as clear over-prescribing of ICS. This is a key finding that could be easily targeted by a quality improvement program reducing both adverse outcomes and costs. I would expect this to be pointed out in the discussion section.

RESPONSE: We agree with this comment. Over-prescription of ICS is a key finding but not extremely new in COPD literature. We have made sure that there is already a paragraph mentioning this in the discussion.

- It is a pity that the proportion of available post-bronchodilator spirometry is not presented because this would really reflect accurate diagnosis. I strongly recommend reporting this, especially since you should have this in your dataset.

RESPONSE: We have now included this information, as per the suggestion.
DISCUSSION

- Page 13 line 22. Again the novelty of the manuscript is emphasized. I believe that novelty is not an important characteristic of a clinical audit. Time passes and medical care changes. Audits don't need to be the first in order to be relevant. Audits should be implemented regularly and inform quality improvement interventions (e.g. in pre-post designs). Besides from being inaccurate the novelty is not important and stating this in the summary of the findings (designated first paragraph in discussion) seems inappropriate to me.

RESPONSE: We have to unfortunately disagree with the reviewer. Highlighting the novelty of a study in the context of current knowledge is one of the goals of a discussion section. After reading all the comments, it seems that the reviewer is particularly disturbed by the fact of us being eventually the first COPD audit in primary care to be published. As we have argued in the response to previous comments, this is true, strictly speaking. We have, however, changed this sentence to a less striking one and have moved it to the second paragraph of the discussion.

- Page 13 line 26: As different auditors have performed the audit, how can we be sure that variability did not originate from auditors? Have you made any interventions or measurements to guarantee the auditors uniformity in extracting the information from the medical charts? If no, this must be stated in the limitation section (information bias).

RESPONSE: Yes, this potential effect is intrinsic to multicenter nationwide clinical audits. It even happens in the UK audits, who have several decades of experience in clinical audits and are far ahead of the rest of European countries. Of note, although this is commented as a potential limitation in the discussion, it is difficult to think that the data would be biased to show a determined effect in a precise direction. To try to compensate for this potential effect, we instructed the investigators on how to complete the dataset, the web page had several rules to ensure the data was correct, and finally, as stated in the statistical section, before performing any analysis, the database was evaluated for quality. The values that were extreme from a clinical point of view, missing, or inconsistent were returned to the investigators for correction. We have completed the methodology with this information.
- Page 13 line 35 to 45: This whole section does not discuss the results of the study (but replicates a section from the Background).

RESPONSE: Again, we kindly disagree with the reviewer. This is a second paragraph that allow us to put into context the following discussion. A similar idea (not a replication) is shown in the introduction, because it is key for the reader to understand the context.

- Page 13 line 47: Strength and limitations should be discussed at the end of the discussion section

RESPONSE: This is an interesting requirement. Although there is no golden rule about this, we normally put the strengths and limitations at the beginning of the discussion, since we understand that this would help the reader put into context the rest of the discussion. We have, however, moved these two paragraphs to the end of the discussion, before the final paragraph, as per the reviewer’s suggestion.

- Page 14 line 43 to 47: In my understanding GOLD 2017 groups (A to D) are most certainly NOT based on pre-specified therapies as the authors state. GOLD groups are specified according to 1-years exacerbations (with or without hospital stay) and impact of disease (CAT or MMRC). The recommendations for pharmacological therapies then apply to the GOLD groups (not the other way).

RESPONSE: The reviewer is correct. We thank for noticing this mistake. We have modified the wording accordingly.

- Page 15 line 22-46: In GOLD A a bronchodilatator therapy is suggested but not strictly recommended by the GOLD guidelines and may also be stopped. Also it must not be a long acting bronchodilatator. I do not agree that the population described here is necessarily under-prescribed.
RESPONSE: We have changed the word under-prescription, as per the suggestion.

- The Discussion section should be re-written and mirror the results. I strongly recommend using a reporting guideline in order to guarantee the structure according to standards of reporting. Also, given that previous results exist (unknown to the authors), the results from this study must be discussed in the light of previous work (not named "audit" but reflecting similar guideline-recommended processes of care) and (optionally) comparing the Spanish healthcare system to the systems where other projects have been carried out.

RESPONSE: We agree that the use of a reporting guideline is always recommended. Unfortunately, as far as we know, there is no ad-hoc reporting guideline available for clinical audits in COPD patients. Without intending to be exhaustive, we have revised the whole discussion to include some comparisons with some of the most recent performance evaluating publications in Primary Care, as per the suggestion.

CONCLUSION

- Page 17 line 17. Again the novelty is stated, see comments above.

RESPONSE: We have provided a detailed explanation in earlier parts of this response letter and would like to request the reviewer to kindly refer to these previous responses.

- Given the methodological uncertainties on where the variation actually comes from, I would leave the unexplained variation out from the conclusion

RESPONSE: With all due respect, we disagree with the reviewer on this point. This conclusion is supported by our data with the limitations already highlighted in the discussion.
- The Conclusion section should be re-written and mirror the main messages from the revised manuscript

RESPONSE: We have revised the conclusion section to ensure all statements are supported by the data.

TABLES

- Table 1: the numbers outside and inside brackets are not understandable with the given labels (e.g. "Male gender (n) 3,159 (73.3)" --> i suppose 73.3 are %. The label should thus say "Male gender n and (%)"). A similar issue arises with Age: (years) cannot possibly be (12.7) on average. This is likely a measure of variation, i suppose the standard deviation. The label should thus say "Age in years (SD)". Please revise the labels to make the reported numbers unambiguously understandable.

RESPONSE: What the values in parentheses represent is explained in the footnote of the table.

- Table 2: under the label of "Final diagnosis correct" now also the symptoms are evaluated. In my understanding symptoms are a measure of impact of the disease and also allow exacerbation identification if any worsening occurs. Again, however, the labeling under "Diagnosis correct" seem inappropriate to me. Please revise or update my understanding if symptoms contribute to correctness of diagnosis. (Again: I believe that symptoms should provoke testing for COPD or treatment of COPD (ideally guided by the CAT Test), but they have no role in actual diagnosis since COPD is even possible without any symptoms.

RESPONSE: Current recommendation documents (mainly GOLD and others) require the presence of previous exposure, symptoms, and obstruction for a correct diagnosis. The previous versions of these documents were not so clear-cut in requiring symptoms for diagnosis. That is why we have left both options in the table.
- Please explain "Inhaler Satisfaction" does this maybe translate to "inhaler technique assessed" or is the inhaler technique a component of "treatment adherence". In summary I seem to miss the inhaler technique assessment which I believe is crucial in COPD care.

RESPONSE: The reviewer is right to point this out. This is not a very strong variable. Since this was a clinical audit, clinicians involved in the steering committee felt that the general satisfaction of the patient with the inhaler should also be evaluated. This was further discussed during various meetings and the selection of this item. This variable does not specifically reflect the correct technique for inhaler use, but if some information indicating that the general satisfaction of the patient with the inhaler was recorded. I understand that this may include different aspects of the inhaler, e.g., comfort of use, ease of use, patient preference, number of inhalations per day and so on. We have included an explanation on this point in the methods section for further clarification.

- Did you include in your set of audited parameters whether smoking status has been determined? As this is the single most important and modifiable prognostic factor I think you should have. Then you should also report this outcome in your audit because doctors should really care about whether their patients do smoke or not and if yes introduce smoking cessation advice and interventions (also not reported in the manuscript [but not audited?]).

RESPONSE: We agree with the reviewer that this is a key relevant aspect of the disease. As referred in the results of the manuscript, smoking status was also evaluated and is included in the manuscript.

- I understand this is hairsplitting but stating to report relative frequencies in brackets, (45.3) actually translates in 45.3 times, not 45.3% (relative frequency 0.453). Please amend by stating "in %" or something similar.

RESPONSE: The meaning of the values in parentheses is explained in the footnote of the table.
APPENDIX

- Are the used audit forms in the appendix?

RESPONSE: The present manuscript does not contain appendices. All relevant information is included in the body of the text.

Heinrich Worth (Reviewer 3)

The authors published data to the important question of the quality of COPD management in primary care. The following remarks have to be made:

i) Did the authors analyze patients with a correct diagnosis with COPD? Did the records include curves for FEV1 and FVC? How many patients labelled as COPD did have no lung function for diagnosing? How many patients did have only prebronchodilation values of COPD? No values of FEV1/FVC were given, They should be included. Did the authors use GLI reference values for lung function? With respect to the data presented it appears that a relevant number of patients under study does not have COPD,

RESPONSE: We thank the reviewer for these comments. We have included all this information in the text or the tables, as per the suggestion. The reviewer is right by pointing out that many of the audited patients did not fulfill the diagnostic criteria for COPD. This is part of the result of the audit, since they were all labelled as having COPD.

ii) Differences in documentation of data and real management may occur. Do the authors have any measures to reduce this problem?

RESPONSE: We agree with the reviewer that this is one key issue in clinical audits, not only to measure, but to provide potential interventions to reduce this problem and improve the procedure. With the current methodology we are in the first step of measuring performance. The next project our group is currently working on is to design an intervention to improve care and then re-audit. Since we do not have that information so far, we have included some aspects for improvement in the discussion.
iii) Were any secondary care interventions during the management of patients labelled as COPD excluded?


iv) Why did the authors group coronary artery disease to vascular conditions instead of grouping the disease into cardiac conditions which has been frequently been done? Was hypertension considered as a comorbidity?

RESPONSE: We thank the reviewer for noticing this. When evaluating the data, we considered that coronary artery disease is a disease of the heart and the vasculature. Accordingly, this was included in both groups. Comorbidities recorded were those being part of the Charlson and COTE indexes. Unfortunately, none of these consider systemic hypertension as a part of the index.

v) Since primary care performance includes management of exacerbations or changes of stable disease it is not convincing that the authors did exclude the management of exacerbations.

RESPONSE: We agree with this limitation of the study. As you do realize, this kind of audit is complex and time-consuming. Therefore, in the steering committee this was thoroughly discussed and it was finally decided to limit the audit to stable cases of COPD. Exacerbations in Primary Care should be assessed in another study. Even with that limitation, this is the best clinical data available on COPD care in Spain.
vi) The authors should explain the meaning of non-reversible obstruction? Does the term include or exclude patients with partially reversible airflow limitation?

RESPONSE: We thank the reviewer for letting us clarify. The current concept of COPD requires a non-reversible airflow obstruction on a spirometry, which means that it does not reverse to normal levels, despite some improvement being observed in some patients.

vii) tab. 3: patient education is missed.

RESPONSE: Yes, patient education was not evaluated in the audit.

Alan Itraja (Reviewer 4):

The paper „Community assessment of COPD Health Care study: a clinical audit on primary care performance variability in COPD care” by Dr. María Abad-Arranz and co-authors has been dedicated to describe a clinical audit performed at the level of primary care to evaluate the clinical care delivered to the patients with COPD in a routine clinical practice. Claimingly, this is the first ever clinical audit-derived analysis carried out in primary care to evaluate the clinical performance on COPD. The manuscript is therefore timely to fill this gap. The diagnosis and management of exacerbations of COPD were, however, left out of the scope this time. The paper describes the full methodology and main results of this audit with regard to guideline adherence in 1) the diagnosis and 2) treatment of COPD in the primary care centers in Spain. Beyond its importance, the paper is well written and is interesting to read. Some concerns still remain.

RESPONSE: We thank the reviewer for the positive comment. We have provided our responses to each one of the queries below.

Major remarks.

1) The Abstract could be re-written to more informative, especially in terms of methods. I.e. providing more exact data for the reader (as the abstract is often the only part that many people
read) on the time frame, selection of the sites, variables to address, and principles how were the analyses made and interpreted. In the results' part of the abstract, it could be mentioned that that most of the evaluated parameters appeared in the range of inadequate performance. Alternatively, this is also a conclusion for this study, as is the large variability in clinical performance across the centers. On the other hand, identification of the determinants of this variability or even the need to do so is probably not the conclusion from this study, but is rather a limitation of it.

RESPONSE: We agree that the abstract is one of the most important parts of the manuscript. Notably, it is limited by the word count. Therefore, sometimes it is impossible to include all data. We have reviewed the abstract and included as much information as possible to reach 350 words, which is the word limit.

2) Interestingly, smoking-related issues were addressed and considerable numbers of active smokers were found, but not performance of the counseling to quit smoking by the health care providers. Why?

RESPONSE: In table 3 we present the number of subjects in whom some recommendation to quit smoking was given.

3) Statistics. The authors state that the significance of variability between the different centers was analyzed by chi-squared test or ANOVA. This is appropriate, if the chi-square test was chi-square test variance. This could be stated in the text. Bartlett's, Levene's or Box's tests would also have been appropriate alternatives for the statistics.

RESPONSE: Thanks for pointing this out. All the statistics were assessed by a statistician to ensure this is correct.

4) Another statistics-related issue is the lack of assessment of the confounding or explanatory variables for the great variability detected and described in this paper. It is
understood that this is a matter of another analysis, however, without an appropriate analysis, the
undersigned is not convinced in that the high variability in the audited parameters is not related
e.g. to FEV1 and the severity of COPD as measured by this parameter.

RESPONSE: We agree with the reviewer on this relevant point. Indeed, variability is key in
clinical practice and one of the key goals of clinical audits, and the reason we are reporting it.
We would like to thank the reviewer for letting us comment. Previous studies have demonstrated
significant variability in the processes and outcomes of COPD care in different settings. It has
been shown that this variability in clinical practice is not exclusively influenced by disease
severity, clinical presentation, or the resources available. In fact, a so-called cluster effect is
described. The cluster effect indicates that patients with similar characteristics may experience
different processes of care and outcomes, depending on the center, because they are subject to
distinct common contextual influences beyond resources or standards of care, like the
characteristics of the catchment area, such as socioeconomic status and utilization of health
services or COPD specific criteria in a determined clinical situation. In our audit, we found an
extremely important variability in clinical practice that should be explored in the future. The idea
behind quality of care is that all patients with the same condition should be managed similarly.
However, this is very complex in clinical practice since clinical practice is influence by a number
of additional factors, including the variation in the clinical presentation of chronic diseases; the
perception of the patient, which is influenced by many personal aspects; and the varied responses
to therapies. Therefore, describing the variability is essential to really understand its potential
implications allowing us to seek for specific approaches in the future. The analysis of
explanatory variables for variability will be a matter of another analysis. We have included a
note in the discussion, as per the suggestion.

5) Results. The text is full of vague formulations like "considerable" (everywhere, but e.g.
in Results, page 11. Last paragraph), "moderate" (e.g. in Results, page 11. Last paragraph),
"minority of" (Results, page 12, second paragraph) etc.
RESPONSE: We agree that these terms imply a subjective evaluation of the data presented. We have used these terms to emphasize some of our findings or aspects of the methodology. We also give the number of all our findings so that the reader can come to their own conclusion.

6) The study has been funded by a pharmaceutical company. This has been mentioned somewhere, but not in the main text. The role of the funding source could be outlined in e.g. the Discussion.

RESPONSE: According to the journal’s instructions, the funding source must be acknowledged at the end of the manuscript in a specific section.

Minor comments.

1) Some references are missing (for example for the GINA consensus document (page 10), the mMRC (page 10) etc.

RESPONSE: Thanks for noticing this. We have added those references.

2) "Ethical considerations" (page 10). Is there a need to go through all the aspects provided? This seems to be rather a matter for an application to obtain ethics clearance. Probably, it is enough to state, that the study protocol was approved by an appropriate ethics review board.

RESPONSE: In a traditional article, this would be the case. However, in the present manuscript we have described the methodology of the audit. Since this is the first audit in our country, we believe that providing these details will help other researchers interested in the same topic. Therefore, we would like to keep it as it is, if possible

3) The resolution of the Figure 3 is poor for assessment.
RESPONSE: We have improved the resolution of figure 3.

4) "Dual bronchodilator" could be used instead of "double b." (page 15, line 48).

RESPONSE: We have to unfortunately disagree with the reviewer. We know that the term dual bronchodilation is extended in the literature. However, this is not strictly correct. Double means to add two different things that are used for the same purpose, whereas dual refers to one drug with two different actions. We request the reviewer to kindly refer to Arch Bronconeumol. 2015 Dec;51(12):661 for a more detailed explanation.

5) Grammar check could be done, e.g. for comma use etc.

RESPONSE: We have sent the manuscript to a native English-speaking editor to check for correctness.