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 ABC Index: Quantifying experienced burden of COPD in a discrete choice experiment and predicting costs
AUTHORS	Goossens, Lucas; Rutten-van Mölken, Maureen; Boland, Melinde; Donkers, Bas; Jonker, Marcel; Slok, Annerika; Salomé, Philippe; van Schayck, Onno; in 't Veen, Johannes; Stolk, Elly

VERSION 1 - REVIEW

REVIEWER	Roberto W. Dal Negro CESFAR - Centro Nazionale Studi di Farmacoeconomia e Farmacoepidemiologia Respiratoria - Verona - Italy
REVIEW RETURNED	01-Jun-2017

GENERAL COMMENTS	The paper assessed the COPD burden mainly from the patients' point of view by means of a novel multidimensional investigational tool (the BFC questionnaire). Authors conclude that this tool is more fitting than already available instruments in easily assessing improvement/deterioration of COPD conditions in clinical practice, and strictly related to the cost of illnesses in those patients. Criticisms: 1 - the ethical approval of the study is not mentioned in the manuscript; 2 - the procedure for calculating the ABC score appears not so easy as claimed, and then reproducibility of results will be affected substantially; 3 - as claimed by Authors, the reproducibility of cost is not homogeneous for patients from different countries; 4 - the major criticisms are due to: a - the absence of any information concerning the presence of patients' comorbidities. In other words, how can the reader be sure that the different ABC scores and related costs reported in the paper are only due to COPD ? What about the relative role of different comorbidities (in particular those related to cardio-vascular, neurological, and metabolic comorbidities) is affecting the results of the study? These aspects should be covered. b - Authors used GOLD guidelines for ranking the patients' severity: so, indices used basically were FEV1 and FEV!/FVC. A much more precise definition of their lung function would be required, just in order to assess and classify the clinical and economic COPD

	burden. The prevalence of emphysema rather then of chronic bronchitis would make a huge difference in results, particularly in patients' perception of the disease, and in costs. This aspect should be implemented in discussion and in the list of study limitations. c - on these bases, Authors' conclusions do appear to cathegoric

REVIEWER	Dr Faye Foster University Campus Oldham, England.
REVIEW RETURNED	05-Jun-2017

GENERAL COMMENTS	This is an astute and well written article which makes its arguments cogently. I enjoyed reviewing it. The only comment I would make is that perhaps you would include a statement that the study obtained ethics approval (or a statement that it was not required), including the name of the ethics committee(s) or institutional review board(s), the number/ID of the approval(s), and a statement that participants gave informed consent before taking part.

REVIEWER	Rune Grønseth Department of Thoracic Medicine, Haukeland University Hospital, Bergen, Norway
REVIEW RETURNED	27-Jun-2017

GENERAL COMMENTS	Referee BMJ-open 2017-017831 Goossens and co-authors report the findings of a discrete choice experiment of COPD patients where they derive weights that are applied to the ABC tool, and subsequently show how the resulting ABC index is able to discriminate groups with different level of healthcare utilization and treatment-related costs. The paper reads well, methods and results are presented clearly. However, the methods could be even more comprehensively described; the analyses could perhaps include more covariates and the discussion is somewhat brief. Abstract and introduction Both the abstract, introduction and methods section emphasize balloon diagrams. However, these are not shown. So perhaps a picture could satisfy the readers curiosity? There seems to be two versions of the second paragraph of the introduction included. Choose one of them. Methods section I can't see any information on ethics approval/written consent. I'd like to see some more information on the respondent selection. How many were approached, how many declined. How many came from primary care centers vs hospital. Do I understand you correctly, when I assume that the index was

	calculated based exclusively on the patients? This is a bit unclear to me. If so, I'd be interested in seeing how the index would look if you did it based on the population-based individuals (as a sensitivity analysis). Please state the cost perspective, and include the actual unit costs. Results If possible, I'd like to see more background information on the participants: smoking habits, number of exacerbations preceding year, chronic respiratory symptoms, drug use et cetera. Also the education level of the patient group seems to have fallen out. Table headings should include more information, to enable the reader to browse them without reading the entire paper. All abbreviations should be listed below the tables. Table 4 lacks units - I guess costs are given in 2016 euros? Discussion I'd like to see more on the external validity based on how participants were selected. Do you have information regarding repeatability of the ABC tool? How does it compare to other tools? If this is given in other publications, it should nevertheless be briefly recapitulated. I'd like to see the difference between the general public and the patients discussed, and related to the existing literature. What are the implications of this difference for economic evaluations, how should utility weights be generated? The problem with this field is that there are a very large number of tools trying to give more comprehensive information regarding the health state of the individual - the BODE index, the GOLD ABCD, the various short forms, EQ-5D, SGRQ, CRQ, CAT and now also the ABC tools. In the discussion I'd like to see the authors view on what this plethora of instruments does to inter-study comparability in COPD studies. What would the minimally clinically important difference be on this 0-100 scale? Appendix There are some language issues, e.g. line 36 (and instead of an) and line 48 (were instead of was).

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Roberto W. Dal Negro
Institution and Country: CESFAR - Centro Nazionale Studi di Farmacoeconomia e Farmaco-epidemiologia Respiratoria - Verona - Italy
Competing Interests: No competing interest

The paper assessed the COPD burden mainly from the patients' point of view by means of a novel multidimensional investigational tool (the BFC questionnaire).

Authors conclude that this tool is more fitting than already available instruments in easily assessing improvement/deterioration of COPD conditions in clinical practice, and strictly related to the cost of illnesses in those patients.

Criticisms:

1 - the ethical approval of the study is not mentioned in the manuscript;

This has been added.

2 - the procedure for calculating the ABC score appears not so easy as claimed, and then reproducibility of results will be affected substantially;

RESPONSE: We have re-written the relevant subsection in the manuscript in order to provide more clarity about the calculations:
'Calculating the ABC Index score
The adjusted regression coefficients from the patient data were used to develop a 0 (best) to 100 (worst) scale. This process consisted of three steps: (1) regression, as described above, (2) linear interpolation of coefficients, (3) re-scaling coefficients.
The original ABC questionnaire presents attributes on a seven-point scale, while the DCE contained only three levels per attribute. For this reason the coefficients of the DCE attribute levels were linearly interpolated, in the second step, in order to create preference weights for all levels in between.
In the third step, the coefficients for the highest levels were combined to represent the unscaled Index score for the worst possible burden of disease. This was then rescaled to a value of 100. Finally, all the coefficients were rescaled accordingly and rounded to the nearest integer. As a result, the ABC Index will always be between 0 and 100.'

3 - as claimed by Authors, the reproducibility of cost is not homogeneous for patients from different countries;

RESPONSE: This is acknowledged in the Discussion section:
'The costing study should be viewed as evidence for the predictive performance of the ABC Index and a first step in a validation process, not as a definitive and precise estimate of healthcare utilisation and costs. Furthermore, given the fact that patients are treated differently in different countries, especially cost estimates only apply to COPD patients in the Netherlands.'

4 - the major criticisms are due to:

a - the absence of any information concerning the presence of patients' comorbidities. In other words, how can the reader be sure that the different ABC scores and related costs reported in the paper are only due to COPD? What about the relative role of different comorbidities (in particular those related to cardio-vascular, neurological, and metabolic comorbidities) is affecting the results of the study? These aspects should be covered.

RESPONSE: We have added comorbidities to table 1, which describes the characteristics of respondents. Considering that we recruited patients from both primary and secondary care, these comorbidities are likely to be representative for the COPD population. We agree that comorbidities can have an impact on patient's well-being and on costs. However, the cost estimates in our study pertain to respiratory disease-related resource use and not total resource use. Similarly, the ABC

questionnaire askes for respiratory symptoms and limitations ‘because of your breathing problems’. The ABC index is based on the average weight that COPD patients attach to various attributes (i.e. the symptoms and limitations). We agree that comorbidities can affect both the scores (=levels) on the attributes as well and that it is conceivable that the weights assigned to these attribute-levels are also affected, when individual experience have an impact on individual preferences. However, this is not likely in this case, given the high degree of concordance of weights derived from patients’ answers and the general public. We have addressed this issue in the revised Discussion section of the paper: ‘It is not certain whether the preference weights would be materially different in other countries or in different samples of COPD patients, but it is conceivable that preferences are linked to individual experience. However, the similarity of the results of patients and members of the general public suggests that the results are robust to selection effects.’

b - Authors used GOLD guidelines for ranking the patients' severity: so, indices used basically were FEV1 and FEV1/FVC. A much more precise definition of their lung function would be required, just in order to assess and classify the clinical and economic COPD burden. The prevalence of emphysema rather than of chronic bronchitis would make a huge difference in results, particularly in patients' perception of the disease, and in costs. This aspect should be implemented in discussion and in the list of study limitations.
c - on these bases, Authors' conclusions do appear to cathegoric

RESPONSE: Besides the GOLD severity categories, we added the average FEV1 value to table 1. Taking lung function into account in the logit model, would have involved estimating very many interaction terms between each of the attribute levels and lung function, because one would need to investigate if these weights are different for different levels of lung function. This would have led to diminished power and, moreover, to interpretation difficulties. We were interested in estimating the average weights that COPD patients assign to different attributes, not in explaining them by relating them to background characteristics, such as lung function. Besides this, the relationship of lung function with symptoms and quality of life is known to be rather weak, as is its relationship with healthcare consumption. For these reasons, we considered it not appropriate to include this information in the analysis.

Reviewer: 2
Reviewer Name: Dr Faye Foster
Institution and Country: University Campus Oldham, England.
Competing Interests: None Declared

This is an astute and well written article which makes its arguments cogently. I enjoyed reviewing it.

The only comment I would make is that perhaps you would include a statement that the study obtained ethics approval (or a statement that it was not required), including the name of the ethics committee(s) or institutional review board(s), the number/ID of the approval(s), and a statement that participants gave informed consent before taking part.

RESPONSE: This statement was added to the Methods section.

Reviewer: 3

Reviewer Name: Rune Grønseth
Institution and Country: Department of Thoracic Medicine, Haukeland University Hospital, Bergen, Norway Competing Interests: None declared

Goossens and co-authors report the findings of a discrete choice experiment of COPD patients where they derive weights that are applied to the ABC tool, and subsequently show how the resulting ABC index is able to discriminate groups with different level of healthcare utilization and treatment-related costs.
The paper reads well, methods and results are presented clearly. However, the methods could be even more comprehensively described; the analyses could perhaps include more covariates and the discussion is somewhat brief.

RESPONSE: We added a more extensive description of the calculation of the Index from the regression results.
Including more covariates in the regression model would, in the case of the logit model, have involved estimating very many interaction terms between the attribute levels and the covariates, because one would need to investigate if the weights are different for different levels of the covariate. This would have led to diminished power and, moreover, to interpretation difficulties. We are interested in estimating the average weights for COPD patients, not in explaining them by relating them to background characteristics.
It would not have been possible to include merely the main effects of additional covariates, since the respondent's characteristics would be the same for both choice alternatives.

Abstract and introduction
Both the abstract, introduction and methods section emphasize balloon diagrams. However, these are not shown. So perhaps a picture could satisfy the readers' curiosity?

RESPONSE: The balloon diagram is included in an earlier publication in this journal (Slok et al. 2016;6) and we are happy to reproduce it here.

There seems to be two versions of the second paragraph of the introduction included. Choose one of them.
RESPONSE: This error has been corrected.

Methods section
I can't see any information on ethics approval/written consent.

RESPONSE: This statement were added to the Methods section.

I'd like to see some more information on the respondent selection. How many were approached, how many declined. How many came from primary care centers vs hospital.

RESPONSE: The numbers of respondents who were approached and who declined are given in the Results section. The proportions of patients from primary care and hospitals was added to table 1.

Do I understand you correctly, when I assume that the index was calculated based exclusively on the patients? This is a bit unclear to me. If so, I'd be interested in seeing how the index would look if you

did it based on the population-based individuals (as a sensitivity analysis).

RESPONSE: Yes, it was based exclusively on the weights that patients assigned to the attributes. This has been clarified in the new version of the subsection 'Calculating the ABC Index score', which was presented earlier in the responses to the first reviewer.
The following table shows the results of the requested sensitivity analysis, when using general public weights. The similarities and differences between patients and general public are not different from those in the regression weights. Both groups largely agree about the importance of the attributes. They attach most weight to fatigue, limitations in moderate physical activities, and also to limitations in social activities, and fear that breathing problems get worse. In general, the public puts relatively more emphasis on symptoms, and, in contrast to the COPD patients, they seemed to assign a greater weight to dyspnoea during exercise than to dyspnoea at rest..

Alternative ABC index scores, based on weights of the general public. Contribution of each domain item to the total ABC index score, per severity level.
None One Two or more
Exacerbations per year 0 5 10

Never Hardly ever A few times Several times Many times A great many times Almost all the time
Fatigue per week 0 1 3 5 7 9 10

Symptoms per week
Dyspnoea at rest 0 1 1 2 3 4 5
Dyspnoea during exercise 0 1 2 3 4 5 7
Coughing 0 0 3 4 4 5 5
Sputum 0 2 3 4 5 6 7

Mental problems per week
Fearing breathing problems 0 1 2 4 5 7 8
Feeling depressed 0 2 4 5 6 7 8
Listlessness 0 1 3 3 3 3 3
Tense feeling 0 0 1 2 3 4 6
Worrying 0 1 2 3 3 3 3

Limitations Not at all Very slightly Slightly Moderately Very Extremely Totally
Strenuous physical activities 0 1 2 3 4 4 5
Moderate physical activities 0 2 3 5 6 8 9
Daily activities 0 1 1 2 4 5 6
Social activities 0 1 2 3 5 6 8

Please state the cost perspective, and include the actual unit costs.

RESPONSE: The perspective has been added to the Methods section, a table with actual unit costs is in the revised Appendix.

Results

If possible, I'd like to see more background information on the participants:
smoking habits, number of exacerbations preceding year, chronic respiratory symptoms, drug use et cetera. Also the education level of the patient group seems to have fallen out.

RESPONSE: We have added smoking habits, number of exacerbations preceding year, comorbidities and education level to table 1.

Table headings should include more information, to enable the reader to browse them without reading the entire paper. All abbreviations should be listed below the tables.

RESPONSE: Table headings were expanded. Abbreviation listings were added below the tables.

Table 4 lacks units - I guess costs are given in 2016 euros?

RESPONSE: They are in 2014 euros. This is stated the table's title.

Discussion
I'd like to see more on the external validity based on how participants were selected.

RESPONSE: The selection criteria were added to the Methods section. A short discussion of the representativeness of the sample was added to the Discussion.

Do you have information regarding repeatability of the ABC tool? How does it compare to other tools? If this is given in other publications, it should nevertheless be briefly recapitulated.

RESPONSE: Information on the ABC tool's validity and reliability was added to the Introduction.

I'd like to see the difference between the general public and the patients discussed, and related to the existing literature. What are the implications of this difference for economic evaluations, how should utility weights be generated?

RESPONSE: The differences and similarities between the preferences of both groups are presented in the Results. Given that we were interested in estimating patient preferences, the ABC Index should be based on their answers, not those of the general public. We are aware of the scientific discussion on whose values should count in the assessment of health-related quality of life outcomes being used in economic evaluations. However, the ABC Index score is not intended to reflect utilities, but to describe the burden of disease as experienced by patients.

The problem with this field is that there are a very large number of tools trying to give more comprehensive information regarding the health state of the individual - the BODE index, the GOLD ABCD, the various short forms, EQ-5D, SGRQ, CRQ, CAT and now also the ABC tools. In the discussion I'd like to see the authors view on what this plethora of instruments does to inter-study comparability in COPD studies.

RESPONSE: The following text was added to the Discussion section:
'The ABC Index is different from severity indicators of COPD, like the BODE-index and the variations

thereof, which include characteristics that were selected based on their ability to predict future outcomes like mortality or disease progression. The ABC index is not a severity indicator but a measure of the burden of disease as experienced by the COPD patient. Compared to health status measures like the SGRQ, the CCQ and the CAT, it is the only instrument in which the weights of the items is entirely based on COPD patients' preferences. Unlike the EQ-5D, which is a generic preference-based instrument, the ABC Index is a disease-specific preference-based instrument. It does not measure utilities for use in cost-utility analyses, like the EQ-5D.'

What would the minimally clinically important difference be on this 0-100 scale?

RESPONSE: That question is difficult to answer. A difference of 1 point in the ABC score can express the difference in one level of one of the attributes. That difference would be noticeable by the patients. In that sense, all such differences would be clinically relevant, if it can be assumed that any noticeable improvement has some value to patients, however small that value may be.
Nevertheless, individual patient's preferences are not necessarily equal to the averages that are used to calculate the ABC index. Whether a patient's burden of disease has improved, is a combination of his changes in the attribute levels and the weights he attaches to the various attributes. Hence, the question whether an individual patient has improved can only be answered by this patient himself. It cannot be based on the ABC Index score, or any other health status instrument, because these instruments might assume weights that do not reflect the preferences of that individual patient.

Appendix
There are some language issues, e.g. line 36 (and instead of an) and line 48 (were instead of was).

RESPONSE: These errors have been corrected.

VERSION 2 – REVIEW

REVIEWER	Rune Grønseth Department of Thoracic Medicine Haukeland University Hospital, Norway
REVIEW RETURNED	15-Aug-2017

GENERAL COMMENTS	Goossens and coauthors have as far as I'm available to judge, provided a satisfactory revision. There are only some sporadic minor errors such as 1) The cost for other hospitalisations lack in the supplementary table 2) Incorrect spelling of "monthly" on page 8