PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	How different are COPD-specific patient reported outcomes, health status, dyspnoea and respiratory symptoms? An observational study in a working population.
AUTHORS	Nishimura, Koichi; Oga, Toru; Nakayasu, Kazuhito; Ogasawara, Miyoko; Hasegawa, Yoshinori; Mitsuma, Satoshi

VERSION 1 – REVIEW

REVIEWER	Susan H Walker Anglia Ruskin University, UK
REVIEW RETURNED	17-Jul-2018

GENERAL COMMENTS	This is an interesting and clearly described study addressing an important topic. I have a few minor comments and suggestions for revision. P.4L.12, P.5L.19 & P7. L.14 the word "conceptions" is used where "conceptual framework" would be more appropriate, as used in P.6L.53. P. 4 L30 where the diagnosis of COPD is described using the "lower limit of normal" - this needs a brief explanation. It is described later in the paper but caused me some confusion when mentioned here. P.6L. 20 'Rasch analysis' perhaps needs a reference

REVIEWER	Liliana Fernandez Trujillo Fundación Valle del Lili, Colombia
REVIEW RETURNED	11-Sep-2018

GENERAL COMMENTS	It is worth bearing in mind that the tools described in the manuscript are designed to evaluate patients with COPD and not healthy individuals. I find interesting the recommendation to extend in future studies the application of similar referrals in non-COPD smokers.

REVIEWER	Kenneth Chapman University of Toronto, Canada
REVIEW RETURNED	13-Jan-2019

GENERAL COMMENTS	The authors have tested 3 instruments used to quantify symptoms of respiratory disease in a large population of workers and report poor correlation amongst the instruments and an inability of any instrument to distinguish COPD from the absence of COPD in smokers.

	Minor comments: 1. The Discussion section of the abstract goes beyond the research findings which should be stated succinctly. 2. One limitation to the study is that patients with asthma could not be excluded reliably by one spirometry. That is, some of the patients regarded as having COPD based on the presence of obstruction may have had asthma. The table lists "history of asthma" and no obvious pattern emerges so these data are reassuring. 3. The data were gathered in a working population which may have narrowed the range of responses. That is, highly symptomatic individuals would likely have retired from the workplace. 4. The study was done in Japanese so that each of the instruments would have been translated from the original language of its development. Although each has been validated in Japanese, it is a limit to the study or to the generalizability of the research.

VERSION 1 – AUTHOR RESPONSE

Reply to the Reviewers

Reviewer number 1

1) P.4L.12, P.5L.19 & P7. L.14 the word "conceptions" is used where "conceptual framework" would be more appropriate, as used in P.6L.53.

Response: According to your advice, the word "conceptions" has been replaced by "conceptual frameworks" in the three parts pointed out.

2) P. 4 L30 where the diagnosis of COPD is described using the "lower limit of normal" - this needs a brief explanation. It is described later in the paper but caused me some confusion when mentioned here.

Response: We agree with your opinion. We are sorry that this part may have been confusing in the original manuscript. Therefore, this sentence in the abstract of the revised manuscript now reads as follows: “In 646 healthy non-smoking subjects, as the reference scores for healthy non-smoking subjects, that is, upper threshold, the Bootstrap 95th percentile values were 1.00 for D-12, 9.88 for CAT and 4.44 for E-RS.”

3) P.6L. 20 'Rasch analysis' perhaps needs a reference

Response: According to your advice, we have added the references here (reference #6 and #7 in the revised manuscript).

Reviewer number 2

1) It is worth bearing in mind that the tools described in the manuscript are designed to evaluate patients with COPD and not healthy individuals. I find interesting the recommendation to extend in future studies the application of similar referrals in non-COPD smokers.

Response: We appreciate your interest. According to your advice, we have modified some sentences in the Discussion section. This part now reads as follows; “The D-12, CAT and E-RS are sufficiently validated for measuring PROs in subjects with COPD, but most participants were not patients with COPD but rather healthy workers. As such, there is a possibility that they are not appropriate tools for the study population. However, since the successful application of the CAT in a working population or a random sampling frame from the populations has also been reported, [19 42] there may be a reason to be hopeful for success with the D-12 and E-RS.” (Page 17, line 18- Page 18, line 3 in the revised manuscript).

Reviewer number 3

1) The Discussion section of the abstract goes beyond the research findings which should be stated succinctly.

Response: We appreciate your suggestion. Since this part has been revised, the discussion section of the abstract in the revised manuscript now reads as follows: “Discussion: Assessment of health status and respiratory symptoms would be preferable to dyspnoea in view of smoking status and airflow limitation in a working population. However, these patient-reported measures were inadequate in differentiating between smokers and subjects with COPD identified by spirometry.”

2) One limitation to the study is that patients with asthma could not be excluded reliably by one spirometry. That is, some of the patients regarded as having COPD based on the presence of obstruction may have had asthma. The table lists "history of asthma" and no obvious pattern emerges so these data are reassuring.

Response: We appreciate your interest. This part of the Results section of the revised manuscript now reads as follows: “There was no difference between groups in the frequency of self-reported history of asthma.” (Page 12, lines 9-10 in the revised manuscript).

3) The data were gathered in a working population which may have narrowed the range of responses. That is, highly symptomatic individuals would likely have retired from the workplace.

Response: We absolutely agree with your opinion. This is one of the reasons why we wrote as follows in the original manuscript; “The D-12, CAT and E-RS are sufficiently validated for measuring PROs in subjects with COPD, but most participants were not patients with COPD but rather healthy workers. As such, there is a possibility that they are not appropriate tools for the study population.”

In addition, we have added the following sentence to the Discussion section: “However, since the successful application of the CAT in a working population or a random sampling frame from the populations has also been reported, there may be a reason to be hopeful for success with the D-12 and E-RS.” (Page 17, line 21- Page 18, line 3 in the revised manuscript).

4) The study was done in Japanese so that each of the instruments would have been translated from the original language of its development. Although each has been validated in Japanese, it is a limit to the study or to the generalizability of the research.

Response: We absolutely agree with your opinion. According to your advice, we have added the following sentences to the Discussion section: “Furthermore, the present study was conducted in Japanese so that each of the instruments would have been translated from the original language of its development. Although the Japanese version has been validated in each case, it may be a limit to the generalizability of the research across the globe.”

VERSION 2 – REVIEW

REVIEWER	Susan Walker Anglia Ruskin University
REVIEW RETURNED	12-Mar-2019

GENERAL COMMENTS	This is an improved version of the paper.

REVIEWER	Kenneth Chapman University of Toronto, Canada
REVIEW RETURNED	07-May-2019

GENERAL COMMENTS	The paper is a simple observational study with clear but unsurprising results. I have two minor thoughts, perhaps for future consideration. 1. Did the paper questionnaire administer the three questionnaires in the same sequence for all subjects? Was there a time interval between questionnaires (for example, to do spirometry)? 2. Was there any review or analysis of discordant results such as those with high dyspnea scores on one instrument and low scores on another? 3. Were items 3 and 4 on the CAT which deal specifically with dyspnea compared to the dyspnea questionnaire?

VERSION 2 – AUTHOR RESPONSE

Reply to the Reviewers

Reviewer number 3

5) Did the paper questionnaire administer the three questionnaires in the same sequence for all subjects? Was there a time interval between questionnaires (for example, to do spirometry)?

Response: We thank you for your questions. Although the specific order of administration of the questionnaires may influence the outcome, we performed the patient-reported outcome assessment in a booklet form, with a fixed order of administration. Randomizing the order of administration should have been performed in this regard. However, there is a lack of scientific evidence suggesting that the specific order of administration could affect the results. For example, several articles comparing questionnaires for asthma did not elaborate on the order of administration (Mölken et al., Eur. Respir. J. 1995; 8: 888-897, Molen et al., Quality of Life Research 1997; 6: 353-361, Molen et al., Eur. Respir. J. 1998; 12: 30 -34). It also means that the time interval between tools is minimal.

Since this part has been revised, the methods section in the revised manuscript now reads as follows: “The Japanese versions of the EXACT, CAT and D-12 were self-administered in the same order under supervision in a booklet form prior to the pulmonary function tests.” (Page 6; lines 16 to 17 in the revised manuscript).

6) Was there any review or analysis of discordant results such as those with high dyspnea scores on one instrument and low scores on another?

Response: Thank you for your question. Accordingly, the following sentences and Table 3 have been added to the results section of the manuscript; “Concordant and discordant results between tools were set to be examined using the above cut-off values (Table 3). However, since there were only a small number of subjects with higher scores on each instrument due to skewed score distribution, those with higher scores on one instrument and lower scores on another were less than one-tenth of all of the

subjects involved." (Page 13; lines 3 to 7 in the revised manuscript).

Table 3. Concordant and discordant results between tools using the cut-off values.

COPD assessment test (CAT) and Evaluating Respiratory Symptoms in COPD (E-RS)

		E-RS Total Score	
		0-4	5 or more
CAT Score	0-9	1,343 (86%)	63 (4%)
	10 or more	113 (7%)	47 (3%)

COPD assessment test (CAT) and Dyspnoea-12 (D-12)

		D-12 Score	
		0-1	2 or more
CAT Score	0-9	1,386 (89%)	20 (1%)
	10 or more	141 (9%)	19 (1%)

Evaluating Respiratory Symptoms in COPD (E-RS) and Dyspnoea-12 (D-12)

		D-12 Score	
		0-1	2 or more
E-RS Total Score	0-4	1,428 (91%)	28 (2%)
	5 or more	99 (6%)	11 (1%)

7) Were items 3 and 4 on the CAT which deal specifically with dyspnea compared to the dyspnea questionnaire?

Response: We appreciate your question as we were also interested in this topic. This issue was addressed in the "Healthcare Professional User Guide. Expert guidance on frequently asked questions. Issue 4: November 2018" where the answer to the question, "Can I just use a few of the questions included in the CAT?" was, "No. The CAT should be used in its entirety. The CAT was validated as an 8-item questionnaire and the questions should not be split up or used independently of each other which will reduce the integrity and measurement properties of the questionnaire." We are therefore hesitant to adopt a similar approach.