

# Comparison of patient-reported outcomes during acute exacerbations of chronic obstructive pulmonary disease

Koichi Nishimura,<sup>1</sup> Saya Nakamura,<sup>2</sup> Masaaki Kusunose,<sup>1</sup> Kazuhito Nakayasu,<sup>3</sup> Ryo Sanda,<sup>1</sup> Yoshinori Hasegawa,<sup>2</sup> Toru Oga<sup>4</sup>

**To cite:** Nishimura K, Nakamura S, Kusunose M, *et al*. Comparison of patient-reported outcomes during acute exacerbations of chronic obstructive pulmonary disease. *BMJ Open Resp Res* 2018;**5**:e000305. doi:10.1136/bmjresp-2018-000305

Received 10 April 2018  
Revised 21 August 2018

## ABSTRACT

**Introduction** The aim of this study was to investigate which patient-reported outcome measure was the best during the recovery phase from severe exacerbation of chronic obstructive pulmonary disease (COPD).

**Methods** The Exacerbations of Chronic Pulmonary Disease Tool (EXACT), the COPD Assessment Test (CAT), the St George's Respiratory Questionnaire (SGRQ), the Dyspnoea-12 (D-12) and the Hyland Scale (global scale) were recorded every week for the first month and at 2 and 3 months in 33 hospitalised subjects with acute exacerbation of COPD (AECOPD).

**Results** On the day of admission (day 1), the internal consistency of the EXACT total score was high (Cronbach's alpha coefficient=0.89). The EXACT total, CAT, SGRQ total and Hyland Scale scores obtained on day 1 appeared to be normally distributed. Neither floor nor ceiling effects were observed for the EXACT total and SGRQ total scores. The EXACT total score improved from 50.5±12.4 to 32.5±14.3, and the CAT score also improved from 24.4±8.5 to 13.5±8.4 during the first 2 weeks, and the effect sizes (ES) of the EXACT total and CAT score were -1.40 and -1.36, respectively. The SGRQ, Hyland Scale and D-12 were less responsive, with ES of -0.59, 0.96 and -0.90, respectively.

**Discussion** The EXACT total and CAT scores are shown to be more responsive measures during the recovery phase from severe exacerbation. Considering the conceptual framework, it is recommended that the EXACT total score may be the best measure during the recovery phase from AECOPD. The reasons for the outstanding responsiveness of the CAT are still unknown.

## INTRODUCTION

Acute exacerbation of chronic obstructive pulmonary disease (AECOPD) is a common cause of emergency hospitalisation and a great risk for morbidity and mortality. Although the definition has been extensively discussed in the literature around the world,<sup>1-4</sup> AECOPD has been clearly defined by a change of the respiratory symptoms in subjects with COPD. The latest Global Initiative for Chronic Obstructive Lung Disease (GOLD) document also definitely describes an exacerbation of COPD as an acute worsening of respiratory symptoms that results in

## Key messages

- ▶ Which patient-reported outcome measure is the best during acute exacerbation of chronic obstructive pulmonary disease (COPD)?
- ▶ Considering the conceptual framework, it is recommended that the Exacerbations of Chronic Pulmonary Disease Tool (EXACT) total score may be the best measure during the recovery phase from exacerbation.
- ▶ This is the first study to draw a direct comparison among five tools, namely the EXACT, the COPD Assessment Test, the St George's Respiratory Questionnaire, the Dyspnoea-12 and the Hyland Scale (global scale), during the recovery phase from severe exacerbation of COPD.

additional therapy.<sup>5</sup> Although an event-based definition of AECOPD has been often used in large-scale clinical trials,<sup>6,7</sup> our concern is possible underestimation of this alternative method to define the AECOPD compared with a symptom-based definition.

How should we make a decision on symptom changes beyond the normal day-to-day variation? The criteria for AECOPD developed by Anthonisen *et al*<sup>8</sup> have historically been the most frequently used method, requiring the occurrence of at least one of the following symptoms: increased dyspnoea, increased sputum volume and increased sputum purulence. Wedzicha and her colleagues have preferred to use the London COPD cohort diary cards, which were developed by their own clinics, in many pioneering research activities.<sup>9-11</sup> Furthermore, studies that have focused on evaluating patient-reported outcomes (PROs) during AECOPD have been reported.<sup>4,12-18</sup> Respiratory-specific questionnaires such as the St George's Respiratory Questionnaire (SGRQ) may be useful to assess the recovery of patients during AECOPD.<sup>19,20</sup> Thus, a standardised method to quantify and evaluate the symptoms needs to be established.



© Author(s) (or their employer(s)) 2018. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>Department of Respiratory Medicine, National Center for Geriatrics and Gerontology, Obu, Japan

<sup>2</sup>Department of Respiratory Medicine, Nagoya University Graduate School of Medicine, Nagoya, Japan

<sup>3</sup>Data Research Section, Kondo P.P. Inc., Osaka, Japan

<sup>4</sup>Department of Respiratory Care and Sleep Control Medicine, Graduate School of Medicine, Kyoto University, Kyoto, Japan

## Correspondence to

Dr. Koichi Nishimura; koichi-nishimura@nifty.com

In light of the growing importance of measuring PROs, some short and simple instruments have become available especially for clinical use since the technological advances in recent years such as the Rasch Analysis have made it possible to reduce the number of items in a questionnaire. Although the COPD Assessment Test (CAT) is a simple instrument originally published by Jones *et al*<sup>21–23</sup> to measure health status in stable COPD, Mackay *et al*<sup>24</sup> evaluated the usefulness of the CAT during AECOPD to assess exacerbation severity, and they concluded that the CAT provides a reliable score of exacerbation severity.

On the other hand, a tool developed by Leidy *et al* that was specifically designed to quantify AECOPD, the Exacerbations of Chronic Pulmonary Disease Tool (EXACT) Patient-Reported Outcome (known as EXACT-PRO), has been reported to be reliable, valid and sensitive to change during exacerbation recovery.<sup>25–27</sup> The EXACT is a diary of recorded symptoms during the AECOPD and conceptually different from health status measurement tools such as the CAT or SGRQ in subjects with stable COPD. Thus, both the EXACT and CAT have been validated to quantify AECOPD.<sup>24 27–29</sup> Therefore, in the present study, we hypothesised that the EXACT would have higher responsiveness than other patient-reported measures in evaluating symptomatic changes during AECOPD.

The aim of this study was to provide an answer to the simple question of which tool is the best for assessing AECOPD and the recovery period in clinical practice. To define the most sensitive measure in detecting the changes during the recovery phase from AECOPD, we investigated and compared the responsiveness of the following outcome measures: the EXACT<sup>25–27</sup> and CAT,<sup>21 24</sup> as well as other patient-reported measures, namely the SGRQ,<sup>19 20</sup> the Dyspnoea-12 (D-12)<sup>30 31</sup> and the Hyland Scale (global scale).<sup>32</sup> The authors investigated and compared the responsiveness of scores obtained from five different PRO measures during the recovery phase from severe exacerbation.

## METHODS

### Study subjects

A total of 33 hospitalised subjects with AECOPD were recruited and followed from the Department of Pulmonary Medicine of the National Center for Geriatrics and Gerontology between September 2013 and March 2016. In the present study, an AECOPD was defined as a worsening of respiratory symptoms that required treatment with oral corticosteroids or antibiotics, or both.<sup>6 7</sup> The inclusion criteria were (1) a clinical diagnosis of COPD, (2) age over 40 years, (3) a history of smoking (10 pack-years or greater), (4) forced expiratory volume in 1 s (FEV<sub>1</sub>)/forced vital capacity (FVC) <0.7 on or before the first day of admission, (5) absence of previous inflammatory changes on chest radiographs that influenced pulmonary function (eg, a previous thoracoplasty or tubercular sequelae), and (6) need to be hospitalised because of the aggravated symptoms of COPD compatible

with exacerbation or severe exacerbation. The exclusion criteria were (1) intubation on the day of admission, (2) comorbidities such as tuberculosis, lung cancer, bronchiectasis and non-tuberculous mycobacteria, (3) non-infectious exacerbations, including exacerbations due to pneumothorax or cardiac failure alone, (4) uncontrolled comorbidity, and (5) disturbed consciousness. Since it is considered that the therapeutic management of AECOPD with no clinical signs of pneumonia should be the same as for an exacerbation of COPD as a result of pneumonia,<sup>33</sup> we included patients with COPD complicated by pneumonia. Although all subjects were evaluated and treated as a general rule according to the international or Japanese guidelines for AECOPD, the type of treatment given was recorded throughout the present study. All eligible subjects with AECOPD underwent pulmonary function tests and arterial blood gas analysis at the baseline and on days 14, 28, 56 and 84 wherever possible. According to the method described by the American Thoracic Society and the European Respiratory Society Task Force in 2005,<sup>34</sup> three acceptable spirometric flow-volume curves were recorded with the patient sitting (Chestac-65V; Chest, Tokyo, Japan). The highest FEV<sub>1</sub> and the highest FVC values among the three manoeuvres were then analysed. The predicted values for the FEV<sub>1</sub> and FVC were calculated according to the proposal from the Japanese Respiratory Society.<sup>35</sup>

### PRO measurement

Participants were requested to complete a self-administered booklet including the Japanese versions of the following five types of the PRO measures at baseline and on days 7, 14, 21, 28, 56 and 84 in the following order: EXACT, CAT, SGRQ, D-12 and Hyland Scale.

The EXACT is a 14-item diary where each attribute or item is assessed on a 5-point or 6-point ordinal scale and summed to yield a total score that is converted to a 0–100 scale, with higher scores indicating a more severe exacerbation.<sup>25–27</sup> Three respiratory symptom domains are embedded in the instrument (breathlessness, cough and sputum, and chest symptoms), and the EXACT total scores are also calculated. The recall period was ‘today’ and patients selected the answer that best described their experience for that day. Although the EXACT should be set to use the last available baseline value to identify a new exacerbation after every event, this system is not applied in the present study since no other PRO measures have similar rules.

Disease-specific health status was assessed with previously validated Japanese versions of the CAT and SGRQ (Japanese version 2).<sup>36 37</sup> The CAT is a questionnaire consisting of eight items scored from 0 to 5 in relation to cough, phlegm, chest tightness, breathlessness going up hills/stairs, activity limitations at home, confidence leaving home, sleep and energy. The CAT scores range from 0 to 40, with a score of 0 indicating no impairment.<sup>21 24 28 37</sup> The SGRQ is a disease-specific instrument

designed to measure impact on overall health, daily life and perceived well-being in patients with obstructive airways disease.<sup>20,36</sup> It consists of 50 items and 76 weighted responses, with scores ranging from 0 to 100. Scores are calculated for three components: symptoms, activity and impact, as well as a total score.

To assess the severity of dyspnoea, we used the Japanese version of the D-12, which consists of 12 items (7 physical items and 5 affective items), each with a 4-point grading scale (0–3).<sup>30,31</sup> The D-12 produces a total score (range 0–36, with higher scores representing more severe breathlessness) and two component scores: physical (items 1–7 with scores ranging from 0 to 21) and emotional (items 8–12 with scores ranging from 0 to 15). Global health was also assessed by the Japanese version of the Hyland Scale, with scores ranging from 0 to 100, where 0='might as well be dead' and 100='perfect quality of life'.<sup>32,38</sup>

The EXACT, CAT, SGRQ, D-12 and Hyland Scale were self-administered under supervision in a booklet form. One of the authors (KoN) reviewed the surveys to ensure that subjects did not unintentionally omit the answers or provide multiple responses to any questions. Although the original developers of the EXACT recommend using an electric version such as that provided on an electronic personal digital assistant (PDA), devices with the Japanese version were not available. Therefore, all the surveys were conducted using a paper-based method without the patients knowing anything about their own previous responses, that is, without informed administration. The completed questionnaires were collected every night during the hospitalisation, and the participants were asked to bring them on their subsequent clinic visits after discharge from a hospital.

Furthermore, since the authors had speculated that both the EXACT and the CAT were worthy of our full attention with respect to the responsiveness in consideration of the literature,<sup>24,28,29</sup> these two questionnaires were given every night during the first 4 weeks. We also intended to make a direct comparison between the two tools based on diary performance.

### Statistical analysis

All results are expressed as mean±SD. Calculating Cronbach's alpha coefficient enabled us to assess the internal consistency. The score distribution of the PRO measures was evaluated by histograms and the Shapiro-Wilk test. The effect size (ES) and the standardised response mean (SRM) were used as responsiveness indexes.<sup>39–43</sup> The former represents the mean change in the score divided by the SD of the baseline scores. The latter represents the mean change in the score divided by the SD of the change in the score. Cohen<sup>39</sup> suggested that ES of 0.2–0.5 were regarded as being small, 0.5–0.8 were moderate and those ≥0.8 were large although the SRM is perhaps the closest to the ES. A p value of less than 0.05 was considered to be statistically significant.

## RESULTS

### Patient characteristics

A total of 33 subjects were enrolled in the present study. They were predominantly male (87.9%), and the average age at initial hospitalisation was 75.3 years (SD=8.8). The mean FEV<sub>1</sub> was 1.02 L at baseline in 31 subjects for whom spirometric measures were available and 1.22 L 3 months later in 30 subjects. Using the classification of severity of airflow limitation of the GOLD criteria,<sup>5</sup> out of 30 patients studied on day 84, 5 subjects (16.6%) were in GOLD 1, 11 (36.7%) in GOLD 2, 8 (26.7%) in GOLD 3 and 6 (20.0%) in GOLD 4. As nine patients (27.2%) were diagnosed with COPD for the first time, they had not received treatment prior to admission. Twenty-four patients (72.7%) were treated with single or dual inhaled bronchodilators. Nineteen subjects (57.6%) were also given inhaled corticosteroids. Five patients (15.2%) were treated with long-term oxygen therapy before admission. Two were receiving continuous positive airway pressure treatment for the comorbidity of obstructive sleep apnoea syndrome.

The average length of stay (LOS) in hospital was 16.5 days, and the median value was 15 days. The LOS was longer than 30 days in two cases (6.1%), both due to complications. All the participants except for one were given varying doses of both oral or intravenous corticosteroids and antibiotics. The relapse of an acute exacerbation was detected during the 84-day study periods in 7 out of 33 subjects (21.2%), and 2 were readmitted to our hospital due to relapsing exacerbation. Three patients could not be followed at the end of the study period since they had moved to other clinics for treatment for unrelated conditions.

### Internal consistency and the distribution of test scores on the day of admission

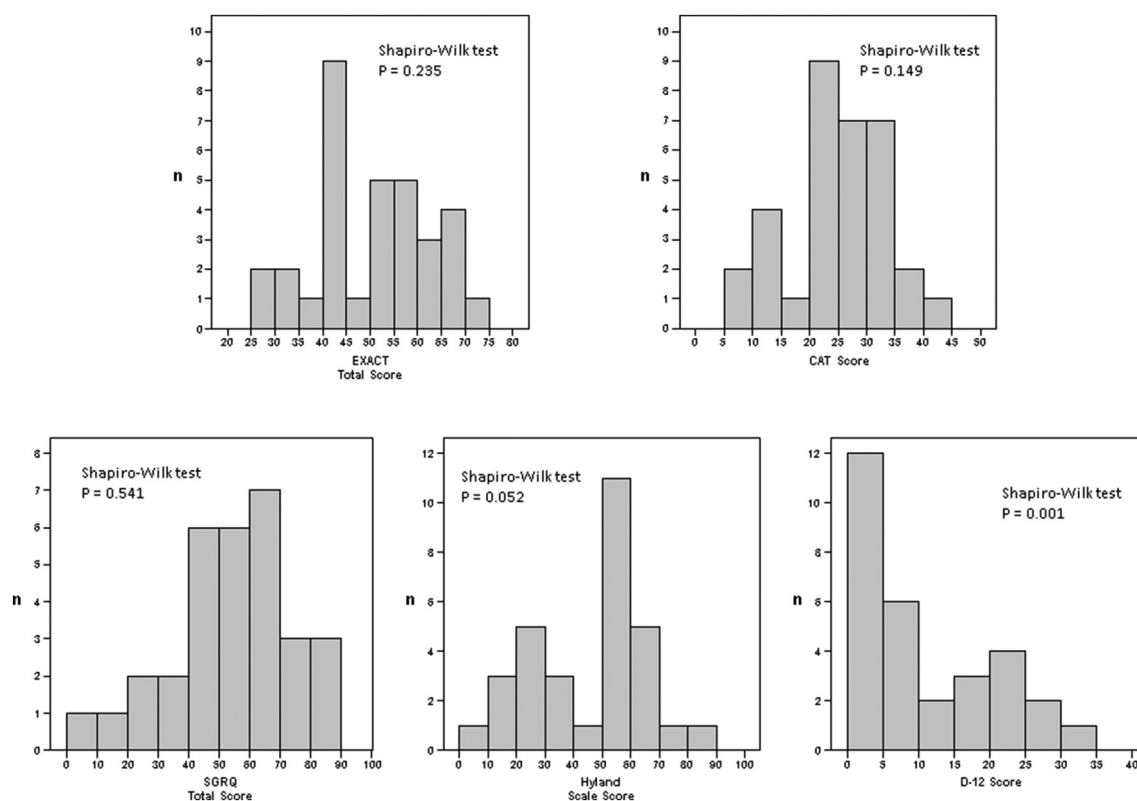
The internal consistency assessed with Cronbach's alpha coefficient was high ( $\alpha=0.89$ ) for the EXACT total score and that of the three domain scores ranged from  $\alpha=0.64$  to  $\alpha=0.93$  (table 1). The internal consistency of the D-12 total score was higher ( $\alpha=0.97$ ) than that of the CAT score and SGRQ total and three components (ranging from  $\alpha=0.78$  to  $\alpha=0.93$ ).

Frequency distribution histograms of the scores obtained are shown in figure 1. Although the EXACT total, CAT, SGRQ total and Hyland Scale scores obtained on day 1 appear to be normally distributed (Shapiro-Wilk tests,  $p=0.235$ ,  $p=0.149$ ,  $p=0.541$  and  $p=0.052$ , respectively), the normality of the score distribution of the D-12 score was rejected using the Shapiro-Wilk tests ( $p=0.001$ ). Neither floor nor ceiling effects were observed for the EXACT total and SGRQ total scores. While a floor effect appeared for the Hyland Scale and D-12, a slight ceiling effect was found for the CAT score (table 1).

**Table 1** Internal consistency and score distribution for each patient-reported outcome obtained on day 1 (on admission day due to acute exacerbation of COPD)

Patient-reported outcomes	Possible score range	Items (n)	Cronbach's alpha coefficient	Score distribution					Floor effect (%)	Ceiling effect (%)
				Mean	SD	Median	Maximum	Minimum		
EXACT total score	0–100	14	0.89	50.5	12.4	52.0	70.0	25.0	0.0	0.0
Breathlessness domain score	0–100	5	0.93	54.7	24.9	56.0	87.0	0.0	6.1	0.0
Cough and sputum domain score	0–100	2	0.65	38.8	21.8	39.0	86.0	0.0	6.1	0.0
Chest symptoms domain score	0–100	3	0.64	48.0	20.0	52.0	79.0	12.0	0.0	0.0
CAT score	0–40	8	0.85	24.4	8.5	25.0	40.0	5.0	0.0	3.0
SGRQ total score	0–100	50	0.93	55.1	19.8	57.0	88.0	6.5	0.0	0.0
SGRQ symptoms	0–100	8	0.78	66.7	20.8	65.8	100.0	6.3	0.0	6.5
SGRQ activity	0–100	16	0.89	67.7	25.9	79.2	100.0	0.0	3.2	6.5
SGRQ impact	0–100	26	0.87	44.2	22.2	45.2	88.0	0.0	3.2	0.0
Hyland Scale score	0–100	1	NA	42.6	20.7	50.0	80.0	0.0	3.2	0.0
D-12 total score	0–36	12	0.97	10.3	9.7	5.0	31.0	0.0	6.7	0.0
D-12 physical score	0–21	7	0.95	7.3	5.9	5.0	20.0	0.0	6.7	0.0
D-12 affective score	0–15	5	0.97	3.2	4.1	1.0	11.0	0.0	48.4	0.0

CAT, COPD Assessment Test; COPD, chronic obstructive pulmonary disease; D-12, Dyspnoea-12; EXACT, Exacerbations of Chronic Pulmonary Disease Tool; SGRQ, StGeorge's Respiratory Questionnaire.



**Figure 1** Frequency distribution histograms of the Exacerbations of Chronic Pulmonary Disease Tool (EXACT) total, the COPD Assessment Test (CAT), the St George's Respiratory Questionnaire (SGRQ) total, Dyspnoea-12 (D-12) total and Hyland Scale scores obtained on day 1 (on admission day due to acute exacerbation of chronic obstructive pulmonary disease (COPD)).



## Responsiveness

The EXACT total score (possible range: 0–100) improved from  $50.5 \pm 12.4$  to  $32.5 \pm 14.3$  and the CAT score (possible range: 0–40) also improved from  $24.4 \pm 8.5$  to  $13.5 \pm 8.4$  during the first 2 weeks (table 2). The speed of the recovery was the fastest during the initial period (figure 2). After that, the EXACT total score made up the transition to  $33.4 \pm 14.9$ , CAT score to  $13.3 \pm 8.4$ , and the SGRQ total score from  $55.1 \pm 19.8$  to  $42.1 \pm 23.3$ , D-12 score from  $10.3 \pm 9.7$  to  $4.3 \pm 5.7$ , and Hyland Scale score from  $42.6 \pm 20.7$  to  $61.0 \pm 21.6$  during the first 28 days (table 2 and figure 3).

The responsiveness of the PRO measures from baseline to 7, 14, 21, 28, 56 and 84 days later was evaluated by ES and SRM (tables 3 and 4). During the first 2 weeks, the ES of the EXACT total and CAT scores were  $-1.40$  and  $-1.36$ , respectively. The SGRQ, Hyland Scale and D-12 were less responsive, with respective ES of  $-0.59$ ,  $0.96$  and  $-0.90$  (table 3). For the same period, the SRMs of the EXACT total and CAT scores were  $-1.13$  and  $-1.32$ , respectively. The SGRQ, Hyland Scale and D-12 were less responsive, with respective SRMs of  $-0.96$ ,  $0.98$  and  $-0.85$  (table 4). Therefore, the EXACT total and CAT scores showed the best responsiveness.

## DISCUSSION

This is the first study to compare the responsiveness among several PRO measures during the recovery phase from severe exacerbation of COPD, and to draw a direct comparison between the details of the EXACT and CAT. First, the present study has achieved outstanding results for the responsiveness of both the EXACT and CAT compared with other PRO measures, including the SGRQ, Hyland Scale as well as the D-12. Second, we failed to identify any significant differences between the EXACT and CAT, although we examined internal consistency and the distribution of the scores on the day exacerbation occurred, magnitude of score change, and ES as well as SRM. Surprisingly, the CAT had a better response than expected despite the core concept that the CAT was developed to measure health status in subjects with stable COPD as opposed to those with AECOPD. Our hypothesis that the EXACT is more responsive than the CAT could not be supported in the present study.

The magnitude of the score change of the PRO measure during AECOPD may depend on many factors, including the severity of the AECOPD, the severity of the baseline COPD, the exacerbation patterns and speeds, and the kind of treatment and management. Andersson *et al*<sup>44</sup> first reported in 2002 that the mean SGRQ total score was 54 units at a severe exacerbation of COPD, and 49 units after 3 and 6 months, a difference of 5 units. Differences in the scores between an exacerbation and the stable period were very large in the Gemifloxacin Long-term Outcomes in Bronchitis Exacerbations (GLOBE) study.<sup>19</sup> In exacerbators who did not relapse, the improvement of the SGRQ total score during the first 4 weeks and

the subsequent 22 weeks was 11.8 and 5.2 units, respectively. The SGRQ total scores and the magnitude of the score improvement after AECOPD in the present study appeared similar to those recorded in previous studies.

On the other hand, the EXACT was developed as a symptom diary designed specifically to quantify exacerbations in COPD. That work conducted by Leidy and a collaborative group including the Food and Drug Administration demonstrated that its 14 items met the criteria for a unidimensional measure of exacerbation severity, providing a daily diary for detecting and quantifying exacerbation severity. Therefore, the EXACT has been used as one of the outcome measures in some large-scale clinical trials.<sup>45 46</sup> Consequently, the EXACT is currently believed to provide valuable information on the course of AECOPD, although some still may hold strong views against it.<sup>29</sup> Although the EXACT was developed to be administered by an electronic device such as a PDA, paper-based questionnaires were used in the present study. This might be related to underestimation of the responsiveness of the EXACT, compared with the CAT.

The speed of recovery after the onset of exacerbation was the fastest between day 1 and day 14 since ES and SRM were the largest for this period in the present study. Although Leidy *et al*<sup>27 46</sup> described the figures showing the mean EXACT total scores on days 1–27 of an exacerbation in their validation study, as well as pooled data from two 12-week phase II international, randomised controlled trials, the pattern of the EXACT total scores appears similar to that in the present study. The scores of the EXACT were likely kept constant after day 14 in both studies.

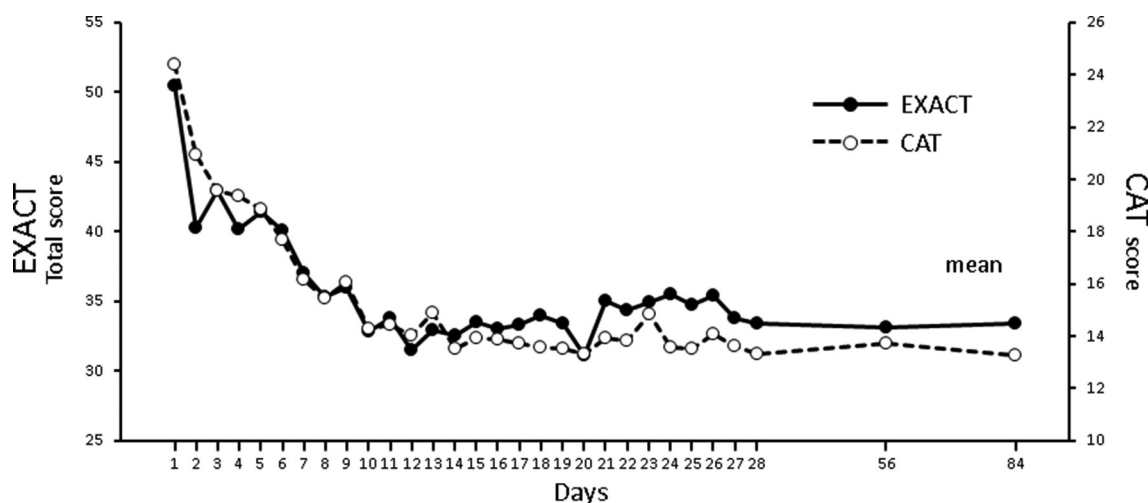
The CAT is a validated simple and short questionnaire designed to assess and quantify health status for subjects with stable COPD. Health status measurements have not been developed to detect changes during short periods, as recall periods are usually over 2 weeks or a month. For example, the most adequate recall period is thought to be 3 months for the SGRQ. Therefore, we were surprised that Mackay *et al* were successful in evaluating the usefulness of the CAT to assess exacerbation severity.<sup>24</sup> In that study, the CAT scores rose from an average baseline value of 19.4 to 24.1 during AECOPD, and the median recovery time to baseline was 12 days. The magnitude and the speed of recovery as reported by the CAT were also similar to the present study but much faster than that reported with the SGRQ in the GLOBE study.<sup>19</sup> The reasons for the outstanding responsiveness of the CAT are still unknown, but one possible explanation may have been due to the mode of the administration as a daily diary. In the present study, since paper-based questionnaires were returned after completion every evening during the hospitalisation, participants completed them without referring to the previous answers, or informed administration. The administration mode may not appear to be related to the high responsiveness of the CAT.

**Table 2** Comparison of each of the patient-reported outcomes and physiological measurement during the recovery phase from acute exacerbation of COPD

Patient-reported outcomes	Day 1		Day 7		Day 14		Day 21		Day 28		Day 56		Day 84	
	n		n		n		n		n		n		n	
Patient-reported outcomes														
EXACT total score	33	50.5±12.4	32	37.0±11.6	31	32.5±14.3	32	35.0±16.2	33	33.4±14.9	30	33.1±14.6	27	33.4±16.0
Breathlessness domain score	33	54.7±24.9	33	34.4±19.4	32	28.3±25.4	32	36.1±26.0	33	33.0±25.0	30	32.9±23.8	27	35.6±26.4
Cough and sputum domain score	33	38.8±21.8	33	26.1±13.9	32	17.2±13.7	32	19.5±14.6	33	19.0±19.2	30	18.4±16.5	27	19.0±15.7
Chest symptoms domain score	33	48.0±20.0	33	27.5±17.2	32	19.4±18.8	32	26.7±20.7	33	21.6±21.0	30	19.8±20.4	27	22.3±21.3
CAT score	33	24.4±8.5	33	16.2±7.3	32	13.5±8.4	32	13.9±8.1	33	13.3±8.4	30	13.7±8.3	27	13.3±8.6
SGRQ total score	31	55.1±19.8	33	49.3±22.4	32	41.8±25.3	31	44.3±23.9	32	42.1±23.3	29	40.3±23.6	26	40.6±25.5
SGRQ symptoms	31	66.7±20.8	33	68.9±17.4	32	63.2±20.7	32	63.5±21.4	32	61.8±23.6	30	56.7±20.1	27	52.4±25.5
SGRQ activity	31	67.7±25.9	33	58.7±28.7	32	51.2±32.5	31	56.5±31.0	32	54.2±30.9	29	51.7±30.9	26	54.8±33.2
SGRQ impact	31	44.2±22.2	33	37.9±24.2	32	29.9±26.4	32	30.0±24.6	33	28.0±22.6	30	29.7±23.3	27	27.3±25.2
Hyland Scale score	31	42.6±20.7	32	59.7±21.2	32	63.8±21.0	30	61.5±21.1	31	61.0±21.6	27	59.8±17.6	26	61.3±20.8
D-12 total score	30	10.3±9.7	30	5.3±5.5	32	3.8±5.3	31	4.6±6.0	32	4.3±5.7	30	3.4±6.0	26	3.8±6.4
D-12 physical score	30	7.3±5.9	32	4.3±3.8	32	2.8±3.6	31	3.4±4.0	32	3.1±3.7	30	2.4±3.6	26	2.8±4.1
D-12 affective score	31	3.2±4.1	31	1.2±2.2	32	1.0±1.9	31	1.3±2.4	32	1.1±2.2	30	1.0±2.6	27	1.0±2.4
Spirometry														
FVC (L)	31	2.17±0.68	33	2.45±0.61	33	2.45±0.61	33	2.45±0.61	33	2.45±0.64	32	2.46±0.65	30	2.46±0.67
FVC (%predicted)	31	71.6±22.8	33	81.0±20.1	33	81.0±20.1	33	81.0±20.1	33	81.6±22.7	32	83.0±23.4	30	81.0±22.3
FEV <sub>1</sub> (L)	31	1.02±0.50	33	1.22±0.53	33	1.22±0.53	33	1.20±0.53	32	1.20±0.53	32	1.19±0.55	30	1.22±0.56
FEV <sub>1</sub> (%predicted)	31	42.7±19.1	33	51.6±20.7	33	51.6±20.7	33	51.3±22.2	32	51.3±22.2	32	51.6±23.8	30	51.3±22.9
Arterial blood gas*														
PaO <sub>2</sub> (mm Hg)	32	67.5±11.4	33	75.6±12.3	33	75.6±12.3	33	76.7±14.4	33	76.7±14.4	33	75.9±11.1	30	73.4±13.4
PaCO <sub>2</sub> (mm Hg)	32	42.0±12.1	33	43.1±8.2	33	43.1±8.2	33	42.7±7.9	33	42.7±7.9	33	42.2±7.8	30	43.7±8.0
pH	32	7.408±0.048	33	7.397±0.030	33	7.397±0.030	33	7.395±0.032	33	7.395±0.032	33	7.401±0.031	30	7.388±0.030

\*Oxygen was administered to six subjects; one subject received non-invasive positive pressure ventilation. Data are presented as mean±SD.

CAT, COPD Assessment Test; COPD, chronic obstructive pulmonary disease; D-12, Dyspnoea-12; EXACT, Exacerbations of Chronic Pulmonary Disease Tool; FEV<sub>1</sub>, forced expiratory volume in 1 s; FVC, forced vital capacity; PaCO<sub>2</sub>, arterial carbon dioxide pressure; PaO<sub>2</sub>, arterial oxygen pressure; SGRQ, St George's Respiratory Questionnaire.

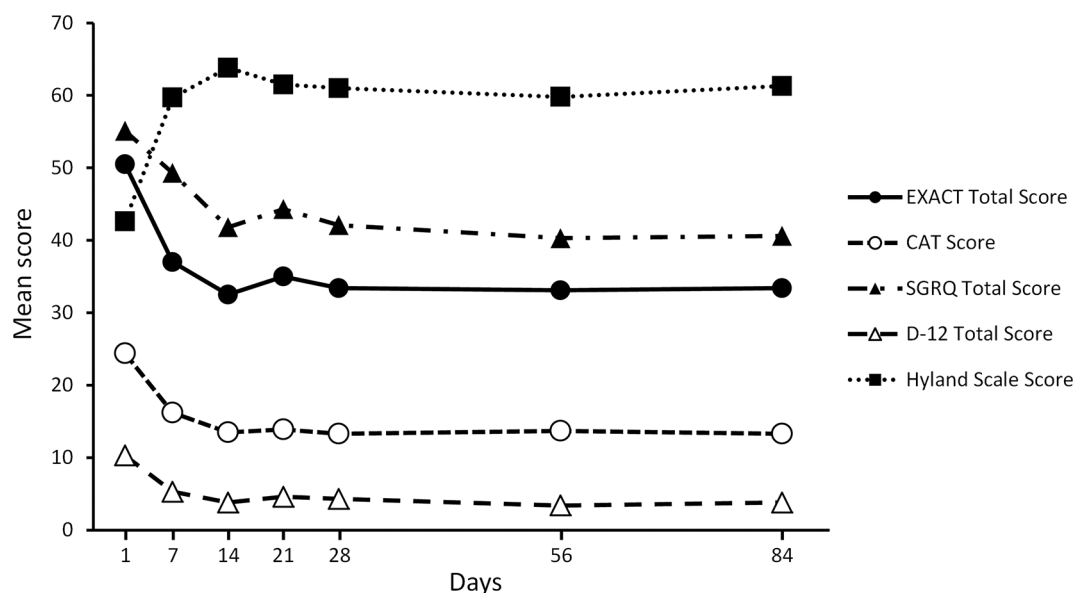


**Figure 2** Comparison of mean daily scores of the Exacerbations of Chronic Pulmonary Disease Tool (EXACT) total and the COPD Assessment Test (CAT) scores over time. Note that possible score ranges of the EXACT total and CAT are 0–100 and 0–40, respectively. COPD, chronic obstructive pulmonary disease.

In the present study, dyspnoea, one of the COPD-specific symptoms, was measured by the D-12 and global perception of health-related quality of life assessed by the Hyland Scale. Although the main reason behind their inclusion was to compare between tools with different concepts, the responsiveness of both instruments was also shown to be high. This may be suggestive of the multicomponent condition of COPD and AECOPD.

We should mention that one of the main limitations of the current evaluation is that only a small proportion of hospitalised patients with AECOPD were included in the present study since there were a number of patients who were unable to provide answers due to

their severe physical incapacity. Another problem was that the present study was limited by the small number of participants included. However, this represents all of the patients with AECOPD who were able to respond adequately to questionnaires in this hospital during the study period. Although it has been reported that the prevalence of COPD in Japan is similar to that in Western countries by a general population sample study, Japanese healthcare providers still feel that COPD is less frequent.<sup>47 48</sup> The LOS in the hospital, which is less than 10 days in most Western countries, is clearly longer in Japan, although this value is about the same as the average LOS for general acute hospitalisations in Japan.<sup>48 49</sup>



**Figure 3** Comparison of mean scores of the Exacerbations of Chronic Pulmonary Disease Tool (EXACT) total, the COPD Assessment Test (CAT), the St George's Respiratory Questionnaire (SGRQ) total, the Dyspnoea-12 (D-12) total and the Hyland Scale scores over time. The possible score ranges of the CAT and D-12 total are 0–40 and 0–36, respectively. Those of the EXACT total, SGRQ total and Hyland Scale scores are all 0–100. COPD, chronic obstructive pulmonary disease.

**Table 3** The effect size for each of the patient-reported outcomes from hospitalisation to 7, 14, 21, 28, 56 and 84 days

Patient-reported outcomes	Days 1-7		Days 1-14		Days 1-21		Days 1-28		Days 1-56		Days 1-84	
	n	Effect size	n	Effect size	n	Effect size	n	Effect size	n	Effect size	n	Effect size
EXACT total score	32	-1.14	31	-1.40	32	-1.09	33	-1.25	30	-1.29	27	-1.18
Breathlessness domain score	33	-0.91	32	-1.08	32	-0.75	33	-0.87	30	-0.89	27	-0.74
Cough and sputum domain score	33	-0.70	32	-1.21	32	-1.03	33	-0.97	30	-1.10	27	-1.14
Chest symptoms domain score	33	-1.10	32	-1.50	32	-1.12	33	-1.29	30	-1.44	27	-1.26
CAT score	33	-1.04	32	-1.36	32	-1.24	33	-1.31	30	-1.28	27	-1.17
SGRQ total score	31	-0.24	30	-0.59	30	-0.47	30	-0.51	27	-0.59	25	-0.52
SGRQ symptoms	31	0.18	30	-0.14	30	-0.05	30	-0.23	28	-0.47	26	-0.38
SGRQ activity	31	-0.32	30	-0.56	30	-0.37	30	-0.37	27	-0.45	25	-0.38
SGRQ impact	31	-0.24	30	-0.59	30	-0.56	31	-0.66	28	-0.61	26	-0.59
Hyland Scale score	31	0.77	30	0.96	28	0.86	30	0.78	26	0.91	25	0.81
D-12 total score	28	-0.54	29	-0.90	28	-0.76	29	-0.76	27	-0.82	24	-0.75
D-12 physical score	29	-0.57	29	-1.04	28	-0.86	29	-0.85	27	-0.99	24	-0.86
D-12 affective score	29	-0.50	30	-0.71	29	-0.60	30	-0.62	28	-0.58	26	-0.61

Data are presented as mean±SD.

CAT, COPD Assessment Test; COPD, chronic obstructive pulmonary disease; D-12, Dyspnoea-12; EXACT, Exacerbations of Chronic Pulmonary Disease Tool; SGRQ, St George's Respiratory Questionnaire.



**Table 4** The SRM for each of the patient-reported outcomes from hospitalisation to 7, 14, 21, 28, 56 and 84 days

Patient-reported outcomes	Days 1–7		Days 1–14		Days 1–21		Days 1–28		Days 1–56		Days 1–84	
	n	SRM	n	SRM	n	SRM	n	SRM	n	SRM	n	SRM
EXACT total score	32	-0.94	31	-1.13	32	-0.92	33	-1.12	30	-1.28	27	-1.10
Breathlessness domain score	33	-0.81	32	-0.96	32	-0.71	33	-0.81	30	-0.95	27	-0.86
Cough and sputum domain score	33	-0.54	32	-0.87	32	-0.83	33	-0.83	30	-0.89	27	-0.84
Chest symptoms domain score	33	-0.79	32	-1.05	32	-0.87	33	-1.05	30	-1.14	27	-0.93
CAT score	33	-0.94	32	-1.32	32	-1.17	33	-1.25	30	-1.21	27	-1.13
SGRQ total score	31	-0.40	30	-0.96	30	-0.71	30	-0.89	27	-0.93	25	-0.72
SGRQ symptoms	31	0.22	30	-0.21	30	-0.06	30	-0.27	28	-0.60	26	-0.39
SGRQ activity	31	-0.41	30	-0.68	30	-0.52	30	-0.51	27	-0.60	25	-0.52
SGRQ impact	31	-0.35	30	-0.86	30	-0.69	31	-0.92	28	-0.82	26	-0.79
Hyland Scale score	31	0.80	30	0.98	28	0.88	30	0.79	26	1.08	25	0.78
D-12 total score	28	-0.59	29	-0.85	28	-0.80	29	-0.82	27	-0.89	24	-0.69
D-12 physical score	29	-0.60	29	-1.01	28	-0.93	29	-0.92	27	-1.13	24	-0.81
D-12 affective score	29	-0.52	30	-0.63	29	-0.54	30	-0.62	28	-0.56	26	-0.51

Data are presented as mean±SD.

CAT, COPD Assessment Test; COPD, chronic obstructive pulmonary disease; D-12, Dyspnoea-12; EXACT, Exacerbations of Chronic Pulmonary Disease Tool; SGRQ, St George's Respiratory Questionnaire; SRM, standardised response mean.

## CONCLUSION

Three main conclusions may be drawn from our findings. First, the EXACT total and CAT scores are shown to be the most responsive measures during the recovery phase from AECOPD among various PRO measures. Second, our hypothesis that the EXACT is more responsive than the CAT could not be supported in the present study since we failed to identify any significant differences between the EXACT and CAT, although we examined the internal consistency and the distribution of the scores on the day exacerbation occurred, magnitude of score change, and ES as well as SRM. Third, surprisingly the CAT had a better response than expected despite the core concept that the CAT was developed to measure health status in subjects with stable COPD as opposed to those with AECOPD. The reasons for the outstanding responsiveness of the CAT are still unknown.

**Acknowledgements** The authors would like to thank Nancy Kline Leidy for permission to use the Japanese version of the EXACT.

**Contributors** KN contributed, as the principal investigator, to the study concept and design, analysis of the results, and writing of the manuscript. SN, MK and RS contributed to performance of the study and acquisition of data. KN contributed to statistical analysis. YH contributed to the interpretation and editing of the manuscript. TO contributed to statistical analysis, and the interpretation and editing of the manuscript. All authors read and approved the final manuscript.

**Funding** This study was partly supported by the Research Funding for Longevity Sciences (27-10) from the National Center for Geriatrics and Gerontology (NCGG), Japan.

**Competing interests** None declared.

**Patient consent** Obtained.

**Ethics approval** Approved by the Institutional Ethics Committee of the National Center for Geriatrics and Gerontology (no 638-4).

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** No additional data are available.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

## REFERENCES

- Rodriguez-Roisin R. Toward a consensus definition for COPD exacerbations. *Chest* 2000;117(5 Suppl 2):398S–401.
- Burge S, Wedzicha JA. COPD exacerbations: definitions and classifications. *Eur Respir J Suppl* 2003;41:46S–53.
- Pauwels R, Calverley P, Buist AS, et al. COPD exacerbations: the importance of a standard definition. *Respir Med* 2004;98:99–107.
- Trappenburg JC, van Deventer AC, Troosters T, et al. The impact of using different symptom-based exacerbation algorithms in patients with COPD. *Eur Respir J* 2011;37:1260–8.
- Vogelmeier CF, Criner GJ, Martinez FJ, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease 2017 report. GOLD Executive Summary. *Am J Respir Crit Care Med* 2017;195:557–82.
- Burge PS, Calverley PM, Jones PW, et al. Randomised, double blind, placebo controlled study of fluticasone propionate in patients with moderate to severe chronic obstructive pulmonary disease: the ISOLDE trial. *BMJ* 2000;320:1297–303.
- Hurst JR, Vestbo J, Anzueto A, et al. Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med* 2010;363:1128–38.
- Anthonisen NR, Manfreda J, Warren CP, et al. Antibiotic therapy in exacerbations of chronic obstructive pulmonary disease. *Ann Intern Med* 1987;106:196–204.
- Donaldson GC, Seemungal TA, Bhowmik A, et al. Relationship between exacerbation frequency and lung function decline in chronic obstructive pulmonary disease. *Thorax* 2002;57:847–52.
- Seemungal TA, Donaldson GC, Bhowmik A, et al. Time course and recovery of exacerbations in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2000;161:1608–13.
- Seemungal TA, Donaldson GC, Paul EA, et al. Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998;157(5 Pt 1):1418–22.

12. Erdal M, Johannessen A, Eagan TM, *et al.* Incidence of utilization- and symptom-defined COPD exacerbations in hospital- and population-recruited patients. *Int J Chron Obstruct Pulmon Dis* 2016;11:2099–108.
13. Koutsokera A, Kiroopoulos TS, Nikoulis DJ, *et al.* Clinical, functional and biochemical changes during recovery from COPD exacerbations. *Respir Med* 2009;103:919–26.
14. Spencer S. Health status measurement in exacerbations of COPD. *Expert Rev Respir Med* 2009;3:573–83.
15. Steer J, Gibson GJ, Bourke SC. Longitudinal change in quality of life following hospitalisation for acute exacerbations of COPD. *BMJ Open Respir Res* 2015;2:e000069.
16. Trigueros Carrero JA. How should we define and classify exacerbations in chronic obstructive pulmonary disease? *Expert Rev Respir Med* 2013;7(2 Suppl):33–41.
17. Xu W, Collet JP, Shapiro S, *et al.* Negative impacts of unreported COPD exacerbations on health-related quality of life at 1 year. *Eur Respir J* 2010;35:1022–30.
18. Nishimura K, Jones PW. Health status and acute exacerbations of chronic obstructive pulmonary disease. In: Giroux JB, Vallee C, eds. *Activities of daily living: performance, impact on life quality and assistance*. New York: Nova Science Publishers Inc, 2013: 1–23.
19. Spencer S, Jones PW. Time course of recovery of health status following an infective exacerbation of chronic bronchitis. *Thorax* 2003;58:589–93.
20. Jones PW, Quirk FH, Baveystock CM, *et al.* A Self-complete Measure of Health Status for Chronic Airflow Limitation: The St. George's Respiratory Questionnaire. *Am Rev Respir Dis* 1992;145:1321–7.
21. Jones PW, Harding G, Berry P, *et al.* Development and first validation of the COPD Assessment Test. *Eur Respir J* 2009;34:648–54.
22. Jones PW, Brusselle G, Dal Negro RW, *et al.* Properties of the COPD assessment test in a cross-sectional European study. *Eur Respir J* 2011;38:29–35. 09031936.00177210 [pii].
23. Jones P, Price D, van der Molen T. Role of clinical questionnaires in optimizing everyday care of chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis* 2011;6:289–96.
24. Mackay AJ, Donaldson GC, Patel AR. Utility of the COPD assessment test (CAT) to evaluate severity of COPD exacerbations. *Am J Respir Crit Care Med* 2012.
25. Jones PW, Chen W-H, Wilcox TK, *et al.* Characterizing and quantifying the symptomatic features of COPD exacerbations. *Chest* 2011;139:1388–94. doi:10.1378/chest.139.4.1388.
26. Leidy NK, Wilcox TK, Jones PW, *et al.* Development of the exacerbations of chronic obstructive pulmonary disease tool (EXACT): a patient-reported outcome (PRO) measure. *Value Health* 2010;13:965–75.
27. Leidy NK, Wilcox TK, Jones PW. Standardizing measurement of chronic obstructive pulmonary disease exacerbations: reliability and validity of a patient-reported diary. *Am J Respir Crit Care Med* 2011.
28. Agustí A, Soler JJ, Molina J, *et al.* Is The CAT questionnaire sensitive to changes in health status in patients with severe copd exacerbations? *COPD* 2012;9:492–8.
29. Mackay AJ, Donaldson GC, Patel ARC, *et al.* Detection and severity grading of COPD exacerbations using the exacerbations of chronic pulmonary disease tool (EXACT). *Eur Respir J* 2014;43:735–44.
30. Yorke J, Moosavi SH, Shuldhham C, *et al.* Quantification of dyspnoea using descriptors: development and initial testing of the dyspnoea-12. *Thorax* 2010;65:21–6.
31. Yorke J, Swigris J, Russell AM, *et al.* Dyspnea-12 is a valid and reliable measure of breathlessness in patients with interstitial lung disease. *Chest* 2011;139:159–64.
32. Hyland ME, Sodergren SC. Development of a new type of global quality of life scale, and comparison of performance and preference for 12 global scales. *Qual Life Res* 1996;5:469–80.
33. Lieberman D, Lieberman D, Gelfer Y, *et al.* Pneumonic vs nonpneumonic acute exacerbations of COPD. *Chest* 2002;122:1264–70.
34. Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005;26:319–38.
35. Society CPFCotJR. The predicted values of pulmonary function testing and arterial blood gas in japanese [in japanese. *Jpn J Thorac Dis* 2001;39.
36. Hajiro T, Nishimura K, Tsukino M, *et al.* Comparison of discriminative properties among disease-specific questionnaires for measuring health-related quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998;157(3):785–90.
37. Tsuda T, Suematsu R, Kamohara K, *et al.* Development of the japanese version of the COPD assessment test. *Respir Investig* 2012;50:34–9.
38. Nishimura K, Oga T, Ikeda A, *et al.* Comparison of health-related quality of life measurements using a single value in patients with asthma and chronic obstructive pulmonary disease. *J Asthma* 2008;45:615–20.
39. Cohen J. *Statistical power analysis for the behavioral sciences*. New York: Academic Press, 1977.
40. Oga T, Nishimura K, Tsukino M, *et al.* Comparison of the responsiveness of different disease-specific health status measures in patients with asthma. *Chest* 2002;122:1228–33.
41. Wright JG, Young NL. A comparison of different indices of responsiveness. *J Clin Epidemiol* 1997;50:239–46.
42. Oga T, Nishimura K, Tsukino M, *et al.* A comparison of the responsiveness of different generic health status measures in patients with asthma. *Qual Life Res* 2003;12:555–63.
43. Fayers PM, Machin D. Multi-Item Scales. *Quality of Life: Assessment, Analysis and Interpretation* 2000:72–90.
44. Andersson I, Johansson K, Larsson S, *et al.* Long-term oxygen therapy and quality of life in elderly patients hospitalised due to severe exacerbation of COPD. A 1 year follow-up study. *Respir Med* 2002;96:944–9.
45. Jones PW, Lamarca R, Chuecos F, *et al.* Characterisation and impact of reported and unreported exacerbations: results from ATTAIN. *Eur Respir J* 2014;44:1156–65.
46. Murray LT, Leidy NK. The short-term impact of symptom-defined COPD exacerbation recovery on health status and lung function. *J COPDFG* 2018;5:27–37.
47. Fukuchi Y, Nishimura M, Ichinose M, *et al.* COPD in Japan: the nippon COPD epidemiology study. *Respirology* 2004;9:458–65.
48. Nishimura K, Yasui M, Nishimura T, *et al.* Clinical pathway for acute exacerbations of chronic obstructive pulmonary disease: method development and five years of experience. *Int J Chron Obstruct Pulmon Dis* 2011;6:365–726.
49. Price LC *et al.* UK National COPD Audit 2003: impact of hospital resources and organisation of care on patient outcome following admission for acute COPD exacerbation. *Thorax* 2006;61:837–42.