Minimum clinically important difference for the COPD Assessment Test: a prospective analysis

Samantha S C Kon, Jane L Canavan, Sarah E Jones, Claire M Nolan, Amy L Clark, Mandy J Dickson, Brigitte M Haselden, Michael I Polkey, William D-C Man

Summary

Background The COPD Assessment Test (CAT) is responsive to change in patients with chronic obstructive pulmonary disease (COPD). However, the minimum clinically important difference (MCID) has not been established. We aimed to identify the MCID for the CAT using anchor-based and distribution-based methods.

Methods We did three studies at two centres in London (UK) between April 1, 2010, and Dec 31, 2012. Study 1 assessed CAT score before and after 8 weeks of outpatient pulmonary rehabilitation in patients with COPD who were able to walk 5 m, and had no contraindication to exercise. Study 2 assessed change in CAT score at discharge and after 3 months in patients admitted to hospital for more than 24 h for acute exacerbation of COPD. Study 3 assessed change in CAT score at baseline and at 12 months in stable outpatients with COPD. We focused on identifying the minimum clinically important improvement in CAT score. The St George's Respiratory Questionnaire (SGRQ) and Chronic Respiratory Questionnaire (CRQ) were measured concurrently as anchors. We used receiver operating characteristic curves, linear regression, and distribution-based methods (half SD, SE of measurement) to estimate the MCID for the CAT; we included only patients with paired CAT scores in the analysis.

Findings In Study 1, 565 of 675 (84%) patients had paired CAT scores. The mean change in CAT score with pulmonary rehabilitation was –2·5 (95% CI –3·0 to –1·9), which correlated significantly with change in SGRQ score (r=0·32; p<0·0001) and CRQ score (r=–0·46; p<0·0001). In Study 2, of 200 patients recruited, 147 (74%) had paired CAT scores. Mean change in CAT score from hospital discharge to 3 months after discharge was –3·0 (95% CI –4·4 to –1·6), which correlated with change in SGRQ score (r=0·47; p<0·0001). In Study 3, of 200 patients recruited, 164 (82%) had paired CAT scores. Although no significant change in CAT score was identified after 12 months (mean 0·6, 95% CI –0·4 to 1·5), change in CAT score correlated significantly with change in SGRQ score (r=0·36; p<0·0001). Linear regression estimated the minimum clinically important improvement for the CAT to range between –1·2 and –2·8 with receiver operating characteristic curves consistently identifying –2 as the MCID. Distribution-based estimates for the MCID ranged from –3·3 to –3·8.

Interpretation The most reliable estimate of the minimum important difference of the CAT is 2 points. This estimate could be useful in the clinical interpretation of CAT data, particularly in response to intervention studies.

Funding Medical Research Council and UK National Institute of Health Research.

Lancet Respir Med 2014; 2: 195–203

Published Online February 4, 2014 http://dx.doi.org/10.1016/S2213-2600(14)70001-3

See Comment page 167

NIHR Respiratory Biomedical Research Unit (S S C Kon MBBS, J L Canavan PhD, S E Jones MSc, C M Nolan BSc, Prof M I Polkey PhD, W D-C Man PhD), **and Harefield Pulmonary Rehabilitation Unit** (C M Nolan, A L Clark BSc, W D-C Man), **Royal Brompton & Harefield NHS Foundation Trust, London, UK; Imperial College, London, UK** (S S C Kon, J L Canavan, S E Jones, C M Nolan, Prof M I Polkey, W D-C Man); **and The Hillingdon Hospital NHS Foundation Trust, Uxbridge, UK** (M J Dickson RGN, B M Haselden PhD)

Correspondence to: Dr Samantha S C Kon, Department of Respiratory Medicine, Harefield Hospital, Harefield UB9 6JH, UK **s.kon@rbht.nhs.uk**

Introduction

The COPD Assessment Test (CAT) is a simple, eight item, health status instrument for patients with chronic obstructive pulmonary disease (COPD), which is highly practical,[1] has good psychometric properties, and has been shown to be responsive to pulmonary rehabilitation and recovery from exacerbation.[2–8] A decrease in CAT score represents an improvement in health status, whereas an increase in CAT score represents a worsening in health status. The CAT has been incorporated into the Global Initiative for Chronic Obstructive Lung Disease (GOLD) combined assessment of COPD as a means of establishing a symptomatic threshold to guide initiation of regular pharmacological treatment.[9] Because it takes only 2–3 min to complete, the CAT has practical advantages over longer-established health status questionnaires such as the St George's Respiratory Questionnaire (SGRQ) and the Chronic Respiratory Questionnaire (CRQ).

The minimum clinically important difference (MCID) is the smallest change in score that patients perceive as beneficial or detrimental, and is useful to aid the clinical interpretation of health status data, particularly in response to intervention. By contrast with other health status questionnaires commonly used in COPD such as the SGRQ and CRQ,[10–12] to our knowledge the MCID of the CAT has not been described.

We aimed to estimate the MCID for the CAT using a range of anchor-based and distribution-based methods in three different clinical scenarios: response to pulmonary rehabilitation, recovery after admission to hospital for acute exacerbation of COPD, and longitudinal change with time. We postulated that the CAT score would improve with pulmonary rehabilitation and recovery from admission to hospital (ie, a decrease in CAT score), but worsen with time in stable patients (ie, an increase in CAT score), and that

change in CAT score would correlate significantly with change in other well-established COPD health status questionnaires.

	Baseline	Change after pulmonary rehabilitation	p value
Mean age (years [SD])	70 (9)	..	..
Sex			
Men	327 (58%)	..	..
Women	238 (42%)	..	..
Mean BMI (kg/m^2 [SD])	27·6 (6·0)	..	..
FEV_1 (% predicted)	47·6 (45·9 to 49·3)	..	..
MRC	3·4 (3·3 to 3·5)	..	..
ISW (m)	210 (199 to 222)	53 (47 to 59)	<0·0001
5STS (s)	15·3 (14·6 to 16·0)	−2·4 (−3·9 to −1·9)	<0·0001
4MGS (m s^{-1})	0·90 (0·88 to 0·93)	0·07 (0·06 to 0·08)	<0·0001
SGRQ			
Total	51·0 (49·3 to 52·6)	−5·0 (−6·1 to −3·8)	<0·0001
Symptoms	64·7 (62·7 to 66·6)	−3·8 (−5·4 to −2·3)	<0·0001
Activities	68·8 (66·7 to 70·8)	−3·7 (−5·3 to −2·1)	<0·0001
Impact	36·5 (34·7 to 38·3)	−5·9 (−7·2 to −4·5)	<0·0001
CRQ			
Total	75·9 (74·1 to 77·7)	14·5 (11·1 to 17·9)	<0·0001
Dyspnoea	13·9 (13·4 to 14·5)	4·7 (3·9 to 5·5)	<0·0001
Fatigue	13·4 (13·0 to 13·8)	3·1 (2·5 to 3·7)	<0·0001
Emotion	30·6 (29·9 to 31·4)	4·2 (3·2 to 5·3)	<0·0001
Mastery	18·0 (17·5 to 18·4)	2·7 (2·1 to 3·4)	<0·0001
CAT	21·4 (20·8 to 22·0)	−2·5 (−3·0 to −1·9)	<0·0001

Data are mean (95% CI) or n (%) unless otherwise specified. BMI=body-mass index. FEV_1=forced expiratory volume in 1 s. MRC=Medical Research Council dyspnoea score. ISW=incremental shuttle walk. 5STS=five-repetition sit-to-stand. 4MGS=4 m gait speed. SGRQ=St George's Respiratory Questionnaire. CRQ=Chronic Respiratory Questionnaire. CAT=COPD Assessment Test.

Table 1: **Baseline characteristics and response to pulmonary rehabilitation (Study 1; n=565)**

	Slope (SE)	y intercept (SE)	*r*	p value
Change in SGRQ				
Total	0·18 (0·03)	−1·53 (0·35)	0·32	<0·0001
Symptoms	0·06 (0·02)	−2·21 (0·34)	0·14	0·0064
Activities	0·10 (0·02)	−2·07 (0·33)	0·25	<0·0001
Impact	0·13 (0·02)	−1·65 (0·35)	0·29	<0·0001
Change in CRQ				
Total	−0·15 (0·01)	−0·29 (0·30)	−0·46	<0·0001
Dyspnoea	−0·34 (0·04)	−0·90 (0·30)	−0·36	<0·0001
Fatigue	−0·46 (0·05)	−1·03 (0·30)	−0·36	<0·0001
Emotion	−0·29 (0·03)	−1·26 (0·28)	−0·37	<0·0001
Mastery	−0·40 (0·05)	−1·39 (0·28)	−0·33	<0·0001
Change in ISW	−0·01 (0·00)	−1·69 (0·35)	−0·14	0·0008
Change in 4MGS	−6·69 (2·17)	−2·03 (0·35)	−0·17	0·0022
Change in 5STS	0·03 (0·03)	−2·39 (0·33)	0·06	0·40

SGRQ=St George's Respiratory Questionnaire. CRQ=Chronic Respiratory Questionnaire. ISW=incremental shuttle walk. 4MGS=4 m gait speed. 5STS=five-repetition sit-to-stand.

Table 2: **Change in COPD Assessment Test score with pulmonary rehabilitation compared with external anchors (Study 1; n=565)**

Methods

Participants

In Study 1, we recruited stable patients with COPD from the Harefield Pulmonary Rehabilitation Unit (Harefield Hospital, London, UK). Inclusion criteria included a diagnosis of COPD, ability to walk 5 m, and no contraindication to exercise.[13] For Study 2, we recruited patients who had an acute exacerbation of COPD diagnosed by a physician and an admission longer than 24 h to acute wards at Hillingdon Hospital (London, UK). In Study 3, we recruited stable patients with COPD from outpatient clinics at Harefield Hospital; patients were deemed stable if they had not had an exacerbation in the 4 weeks before the baseline measurement. Patients were not excluded from Study 3 if they had an excerbation during the 12 month follow-up period. Age 35 years or older was an inclusion criterion for all studies. Patients unable to read or understand English were excluded from all studies. All participants gave written informed consent and all studies received local ethics committee approval.

Procedures

Study 1 assessed response of the CAT to pulmonary rehabilitation. We measured the CAT, SGRQ, and CRQ before and after an 8 week outpatient pulmonary rehabilitation programme, consisting of twice weekly supervised exercise and education sessions.[6,12,14] We measured the incremental shuttle walk, five-repetition sit-to-stand, and 4 m gait speed to assess change in physical performance.[15–17] Patients were kept masked to their performance and were asked to rate their change in health status after pulmonary rehabilitation using an adapted five point Global Rating of Change Questionnaire (GRCQ).[18] Patients were asked to classify how they felt after pulmonary rehabilitation according to five responses: "1: much better"; "2: a little better"; "3: no change"; "4: a little worse" and "5: much worse". Data from 255 of the 565 patients in Study 1 have been used as a comparator group in a previous study.[19]

Study 2 assessed change in CAT score after admission to hospital for acute exacerbation of COPD. Forced expiratory volume in one second (FEV_1), CAT, SGRQ, and 4 m gait speed were measured on the day of hospital discharge and about 3 months later.

Study 3 assessed longitudinal change in CAT score with time. Patients with COPD attending outpatient respiratory clinics were asked to complete FEV_1, incremental shuttle walk, 4 m gait speed, CAT, and SGRQ, and again at 12 months.

Statistical analysis

Study 1 was a pragmatic observational study in which we aimed to prospectively recruit a minimum of 500 patients with paired CAT measurements. We anticipated 25% dropout on the basis of our experience with the pulmonary rehabilitation programme and therefore

recruited 675 patients. For studies 2 and 3, 124 paired measurements were required to detect a correlation coefficient above 0·3 between change in CAT and change in anchor with 80% power at the 0·01 significance level. We anticipated a minimum dropout of 30% on the basis of data from other longitudinal cohorts, therefore we recruited 200 patients in each study.

We did data analyses and graphical presentations using SPSS (version 21) and Prism (version 5). When estimating MCID in change in CAT score, only data from participants who completed paired CAT measurements were included for analysis. We used the paired *t* test to compare paired measurements. We used Pearson's correlation (in which the null hypothesis was defined as no correlation) and linear regression to compare change in CAT score with other outcome measurements.

Estimation of MCID

For anchor-based estimation of MCID, predefined criteria for establishing the validity of external anchors were: a significant correlation between change in CAT score and change in anchor, and a correlation coefficient of more than 0·3 as previously recommended.[20] In Study 1, change in CAT score with pulmonary rehabilitation was anchored against change in SGRQ total score, CRQ total score, and CRQ domain scores. For the GRCQ, we calculated the mean (95% CI) change in CAT score with pulmonary rehabilitation in those reporting feeling "a little better". We did not include those reporting feeling "much better" because we believed that including these patients might lead to an overestimation of the MCID. In studies 2 and 3, change in CAT score was anchored against change in SGRQ.

In studies 1 and 2, the focus was to establish the minimum clinically important improvement because of the small numbers of patients reporting significant worsening of health status. For MCID at the individual patient level, we used receiver operating characteristic curves. The change in CAT score cutoff that best discriminated between patients who improved their health status by the established MCID in the SGRQ total score (–4 point change) or CRQ total score (+10 point change) was defined as the MCID, with equal weighting given to sensitivity and specificity.[10,12] For MCID at the population level, we used linear regression analysis to estimate change in CAT score corresponding to the minimum clinically important improvement for the SGRQ and CRQ total scores, and CRQ domain scores (+2·5 dyspnoea, +2·0 fatigue, +3·5 emotion, +2·0 mastery).[10] In Study 3, because there were much the same numbers of patients showing improvement and worsening of health status, we applied receiver operating characteristic curves and linear regression to investigate both minimum clinically important improvement and deterioration.

For distribution-based methods, we calculated half the SD (0·5 SD)[21] and the SE of measurement (SEM),[22] given by the equation: $SEM = SD \times \sqrt{(1-[\text{test-retest reliability}])}$. On the basis of previous data, we assumed the test-retest reliability of the CAT to be 0·8.[2,4]

Role of the funding source

The sponsors of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. All authors had full access to all the data in the study. WD-CM made the final decision to submit for publication.

Results

Study 1 took place between April 1, 2010, and Dec 31, 2012. Of 675 patients with COPD referred for pulmonary rehabilitation, 565 (84%) completed pulmonary

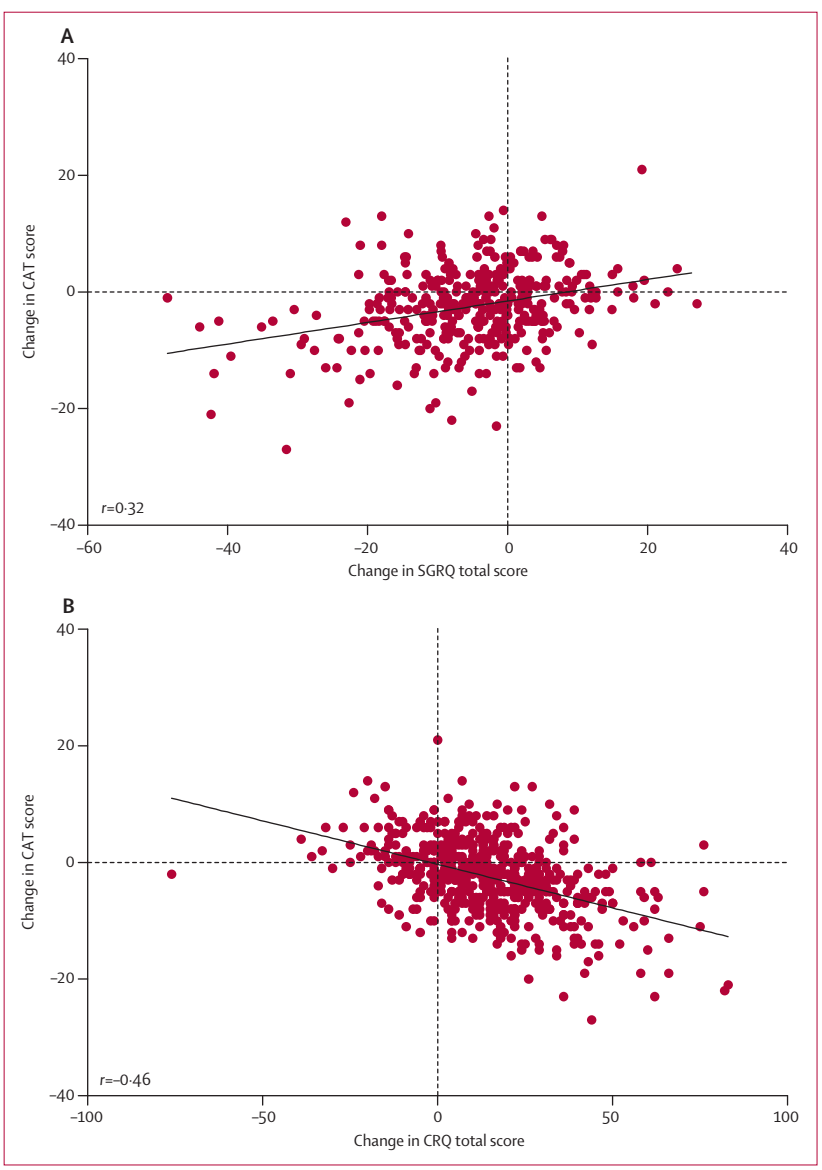

Figure 1: **Association between change in CAT score and change in (A) SGRQ score and (B) CRQ score with pulmonary rehabilitation (Study 1)**
CAT=COPD Assessment Test. SGRQ=St George's Respiratory Questionnaire. CRQ=Chronic Respiratory Questionnaire.

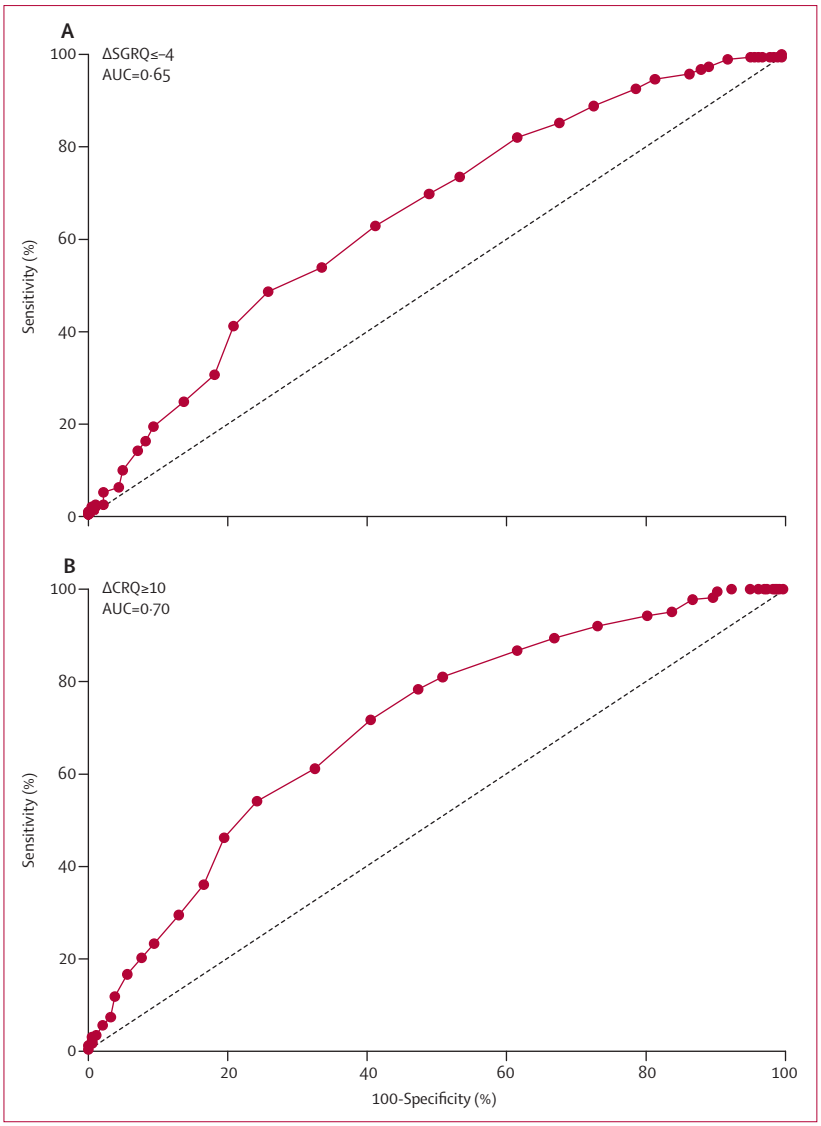

Figure 2: **Receiver operator characteristic curves using a change in COPD Assessment Test score of –2 to best predict achievement of minimum clinically important improvement in (A) SGRQ score and (B) CRQ score (Study 1)**

ΔSGRQ=change in St George's Respiratory Questionnaire score. ΔCRQ=change in Chronic Respiratory Questionnaire score. AUC=area under the curve *C* statistic.

	At hospital discharge	Change at 90 days	p value
Mean age (years [SD])	71 (11)	..	..
Sex			
Men	88 (60%)	..	..
Women	59 (40%)	..	..
Mean BMI (kg/m^2 [SD])	27·1 (5·2)	..	..
FEV_1 (% predicted)	42 (39 to 46)	5·7 (1·7 to 9·8)	0·19
MRC	3·9 (3·7 to 4·1)	–0·6 (–0·8 to –0·4)	<0·0001
SGRQ			
Total	57·1 (54·6 to 59·6)	–6·8 (–8·8 to –4·9)	<0·0001
Symptoms	67·3 (64·2 to 70·3)	–0·3 (–2·9 to 2·3)	0·84
Activities	74·3 (70·9 to 77·7)	–6·0 (–8·8 to –3·1)	<0·0001
Impact	44·1 (41·4 to 46·8)	–9·5 (–11·8 to –7·2)	<0·0001
4MGS (m s^{-1})	0·66 (0·62 to 0·70)	0·22 (0·18 to 0·26)	<0·0001
CAT	23·5 (22·3 to 24·8)	–3·0 (–4·4 to –1·6)	<0·0001

Data are mean (95% CI) or n (%) unless otherwise specified. BMI=body-mass index. FEV_1=Forced expiratory volume in 1 s. MRC=Medical Research Council dyspnoea score. SGRQ=St George's Respiratory Questionnaire. 4MGS=4 m gait speed. CAT=COPD Assessment Test.

***Table 3*: Baseline characteristics (at hospital discharge) of patients with acute exacerbation of COPD and change over 3 months (Study 2; n=147)**

rehabilitation and had paired CAT results. The reasons for non-completion were exacerbation or admission to hospital (46 patients, 7%), poor adherence to pulmonary rehabilitation (29, 4%), work reasons (11, 2%), family illness (ten, 1%), holiday (seven, 1%), declined end of course assessment (four, 1%), and death (three, <1%). Baseline characteristics of the non-completers are shown in the appendix; patients who did not complete pulmonary rehabilitation had significantly worse baseline health status as determined by their SGRQ Impact and CRQ total scores. Table 1 shows the baseline characteristics of the completers, and the changes in health status and physical performance with pulmonary rehabilitation. Mean change in CAT with pulmonary rehabilitation was –2·5 (95% CI –3·0 to –1·9).

See Online for appendix

Table 2 shows the slope, y intercept, and correlation coefficient between change in CAT score and change in other outcome measures. Only change in SGRQ total and change in CRQ total and CRQ domain scores correlated with change in CAT score with a correlation coefficient greater than 0·3 (figure 1). Figure 2 shows the receiver operating characteristic curves for change in CAT in identifying patients who improved their SGRQ and CRQ total scores by more than the established MCID. Both curves were consistent in identifying –2 as the best discriminating CAT change cutoff with an area under the curve *C* statistic of 0·65 for SGRQ and 0·70 for CRQ. The appendix shows the sensitivity and specificity for alternative estimates of the minimum clinically important improvement in CAT score.

With linear regression, by use of a change in CRQ total score of +10 as the cutoff for minimum clinically important improvement, the estimated minimum clinically important improvement in CAT score was –1·8 (95% CI –2·6 to –1·0). By use of the established MCID for the CRQ domains as cutoffs, the estimates for minimum clinically important improvement in CAT score were –1·7 (95% CI –2·5 to –1·0) for dyspnoea, –2·0 (–2·7 to –1·2) for fatigue, –2·3 (–3·0 to –1·5) for emotion, and –2·2 (–2·9 to –1·5) for mastery. With SGRQ total change of –4 as the cutoff, the estimated minimum clinically important improvement for CAT score was –2·3 (95% CI –2·7 to –1·8). There were few patients reporting significant worsening in health status

(figure 1); comparison of the slopes of CAT score against CRQ and SGRQ scores in patients who improved and those who deteriorated showed no significant difference (ANCOVA p=0·67 and p=0·99, respectively).

With the GRCQ, 299 of 565 (53%) patients reported feeling much better, 209 (37%) patients reported feeling a little better, 40 (7%) patients reported no change, and 17 (3%) patients reported feeling a little or much worse. The mean change in CAT score in those reporting feeling "a little better" after pulmonary rehabilitation was –1·6 (95% CI –2·6 to –0·8). For those feeling "much better", the mean change in CAT score was –3·2 (95% CI –4·0 to –2·5).

With distribution-based methods, the estimate for significant improvement in CAT score using 0·5 SD was –3·8 and using SEM was –3·3.

Study 2 took place between Nov 16, 2011, and Dec 31, 2012. Of 200 patients with acute exacerbations of COPD recruited, 147 (74%) had paired CAT measurements. Reasons for dropout were death (14 patients, 7%), declined attendance (26, 13%), acutely unwell in hospital (ten, 5%), and moved out of area (three, 2%). The mean duration between baseline (discharge from hospital) and follow-up measurement was 90·2 days (SD 7·7; range 79–102). The appendix shows baseline characteristics of patients who did not attend follow-up; baseline characteristics of those who attended follow-up and those who did not attend did not differ. Table 3 shows clinical characteristics of this group and the change in outcomes after hospital discharge. Mean CAT score at hospital discharge (23·5 [95% CI 22·3–24·8]) was significantly higher than baseline CAT score in patients referred for pulmonary rehabilitation in Study 1 (21·4 [20·8 to 22·0]) and stable outpatients in Study 3 (20·1 [19·1 to 21·2]; ANOVA p=0·0002).

Mean change in CAT score from hospital discharge to 3 months after discharge was –3·0 (95% CI –4·4 to –1·6). Change in CAT score correlated significantly with change in SGRQ (r=0·47; p<0·0001; figure 3) and change in FEV_1 (r=–0·26; p=0·0021), but not change in 4 m gait speed (r=–0·10; p=0·24). By use of receiver operating characteristic curves, a –2 change in CAT score best discriminated patients who improved by the MCID or more in SGRQ (n=84) with an area under the curve *C* statistic of 0·66. Sensitivity and specificity data for alternative estimates are in the appendix. Linear regression analysis, using an SGRQ change of –4 as the cutoff, estimated the minimum clinically important improvement of the CAT as –2·8 (95% CI –3·7 to –1·9). The slopes of CAT against SGRQ in patients who improved and in those who deteriorated were not significantly different (ANCOVA p=0·21).

With distribution-based methods, the estimate for significant improvement in CAT score using 0·5 SD was –3·7 and using SEM was –3·3.

Study 3 recruited between Jan 1, 2012, and Aug 31, 2012. Of 200 stable patients recruited, 164 (82%) returned for measurements 12 months later. Reasons for dropout included: death (six patients, 3%), current admission to hospital for exacerbation (seven, 4%), current admission to hospital for another reason (four, 2%), declined (ten, 5%), unable to contact (six, 3%), and moved out of area (three, 2%). The mean duration between baseline and follow-up measurement was 364·6 days (SD 20·7, range 332–401 days). The appendix shows baseline characteristics of patients who did not attend follow-up; patients who did not attend follow-up were younger and had better CAT scores than those who attended follow-up. Table 4 shows clinical characteristics of this cohort and change in FEV_1, incremental shuttle walk, 4 m gait speed, CAT, and SGRQ with time. There was no significant change in CAT over time, but change in CAT correlated significantly with change in SGRQ (r=0·36; p<0·0001; figure 4),

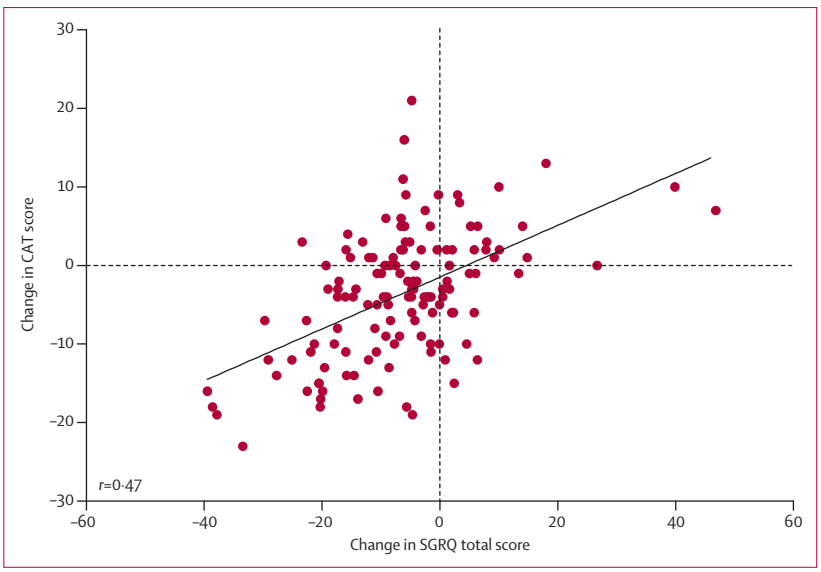

Figure 3: **Association between CAT score and change in SGRQ score at hospital discharge for acute exacerbation of COPD to 90 days after discharge (Study 2)**
CAT=COPD Assessment Test. SGRQ=St George's Respiratory Questionnaire.

	Baseline	Change over 12 months	p value
Mean age (years [SD])	70 (8)	..	..
FEV_1 (% predicted)	47·6 (44·4 to 50·8)	–1·7 (–6·9 to 3·4)	0·51
MRC	3·1 (3·0 to 3·3)	0·1 (–0·1 to 0·2)	0·35
ISW (m)	227 (208 to 247)	–15 (–43 to 12)	0·28
4MGS (m s^{-1})	0·92 (0·89 to 0·95)	–0·04 (–0·06 to –0·02)	<0·0001
SGRQ			
Total	50·6 (48·0 to 53·1)	–0·3 (–4·4 to 3·9)	0·91
Symptoms	63·9 (60·9 to 67·0)	–2·5 (–7·4 to 2·4)	0·32
Activities	68·7 (65·4 to 72·0)	1·3 (–3·8 to 6·3)	0·62
Impact	35·9 (33·1 to 38·7)	0·0 (–4·3 to 4·3)	0·99
CAT	20·1 (19·1 to 21·2)	0·6 (–0·4 to 1·5)	0·25

Data are mean (95% CI) unless otherwise specified. FEV_1=forced expiratory volume in 1 s. MRC=Medical Research Council dyspnoea score. ISW=incremental shuttle walk. SGRQ=St George's Respiratory Questionnaire. 4MGS=4 m gait speed. CAT=COPD Assessment Test.

Table 4: **Baseline characteristics in patients with stable COPD and change over 12 months (Study 3; n=164)**

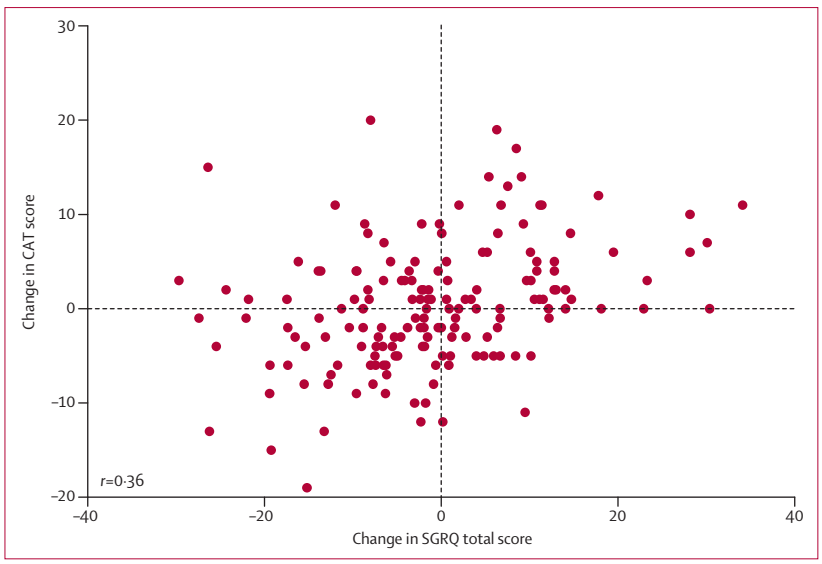

Figure 4: **Association between change in CAT score and change in SGRQ score over 12 months in stable patients with COPD (Study 3)**
CAT=COPD Assessment Test. SGRQ=St George's Respiratory Questionnaire.

incremental shuttle walk (r=–0·25; p=0·0005), and 4 m gait speed (r=–0·18; p=0·0142). For those who improved their SGRQ (n=64), the receiver operating characteristic curve showed that a –1 point change in CAT score was the best discriminant with an area under the curve *C* statistic of 0·66, and that a +1 point change in CAT score best identified those who had a clinically significant worsening in health status as measured by the SGRQ (n=51) with an area under the curve *C* statistic of 0·69. With SGRQ change of –4 as an anchor, linear regression analysis estimated the minimum important improvement in CAT score as –1·2 (95% CI –2·5 to –0·0). The slopes of CAT score against SGRQ score in patients who improved and those who deteriorated were not significantly different (ANCOVA p=0·20).

With distribution-based methods, the estimate for significant improvement in the CAT score using 0·5 SD was 3·8 and using SEM was 3·4. Table 5 summarises all estimates for the MCID for the CAT.

Discussion

Our studies show that the CAT is responsive to change after pulmonary rehabilitation and recovery from hospital admission for acute exacerbation of COPD, and that change in CAT correlates significantly with change in other health status measures and physical performance measures. Furthermore, as far as we are aware, these studies are the first to prospectively and purposively estimate the MCID for the CAT. With anchors measuring similar constructs in different cohorts, the estimates for the minimum important improvement in CAT score ranged from –1·2 to –2·8 with –2 being the most consistent estimate from sensitivity and specificity analyses.

	Anchor or method	MCID
Study 1: Pulmonary rehabilitation		
ROC	SGRQ	–2
ROC	CRQ	–2
Linear regression	SGRQ	–2·3
Linear regression	CRQ	–1·8
Linear regression	CRQ-D	–1·7
Linear regression	CRQ-F	–2·0
Linear regression	CRQ-E	–2·3
Linear regression	CRQ-M	–2·2
Linear regression	GRCQ*	–1·6
Distribution	0·5 SD	–3·8
Distribution	SEM	–3·3
Study 2: At discharge from hospital		
ROC	SGRQ	–2
Linear regression	SGRQ	–2·8
Distribution	0·5 SD	–3·7
Distribution	SEM	–3·3
Study 3: Longitudinal		
ROC	SGRQ	–1
Linear regression	SGRQ	–1·2
Distribution	0·5 SD	–3·8
Distribution	SEM	–3·4

MCID=minimum clinically important difference. ROC=receiver operating characteristic curves. SGRQ=St George's Respiratory Questionnaire. CRQ=Chronic Respiratory Questionnaire total score. CRQ-D=dyspnoea domain. CRQ-F=fatigue domain. CRQ-E=emotion domain. CRQ-M=mastery domain. GRCQ=Global Rating of Change Questionnaire. 0·5 SD=half SD. SEM=SE of measurement. *In patients feeling "a little better".

Table 5: **Anchor-based and distribution-based estimates of the MCID of the CAT in the three studies**

Health status is recommended as an essential outcome measure by pulmonary rehabilitation guidelines.[23] As far as we are aware, Study 1 was the largest study so far to use the CAT during pulmonary rehabilitation (panel). We showed longitudinal validity of the CAT by identifying significant correlations between change in CAT and change in SGRQ, CRQ, and physical performance measures. We recorded a mean change in CAT score of –2·5 with an 8 week outpatient pulmonary rehabilitation programme, in line with previous studies from the UK and the USA and Canada.[3,7] The CAT has also been studied in an unselected chronic respiratory disease population undergoing pulmonary rehabilitation, with 110 non-COPD patients showing a mean CAT change of –2·1, much the same as the mean change measured in patients with COPD.[19]

Several studies have used the CAT as a measure of health status during hospital and community-based treatment of acute exacerbations of COPD, showing changes in CAT ranging from –1·4 to –9·9.[4,5] Study 2 focused exclusively on recovery of the CAT score after admission to hospital, with the baseline CAT measured on the day of hospital discharge. The recorded mean

change in CAT was –3·0, and this outcome correlated significantly with change in SGRQ (r=0·47). This finding could have implications when designing clinical intervention trials in the period after admission to hospital, particularly with regard to sample size calculation.

Until now, there have been few data regarding longitudinal change in CAT with time. In the original description of the CAT, test-retest reproducibility over 7 days was good (intraclass correlation coefficient 0·8), while in stable patients over 4 weeks, the CAT showed a test-retest intraclass correlation coefficient of 0·83.[2,4] Over a 6 month period, Dodd and colleagues[6] showed an increase in CAT from 19·2 to 20·7, although importantly, baseline measurements were recorded immediately after completion of a pulmonary rehabilitation programme. We recorded paired measurements 12 months apart. Our prespecified hypothesis that health status would deteriorate significantly over 12 months (ie, increase in CAT score) was not supported by our data, although deteriorating health status, including death, was a reason for loss to follow-up. Although the 12 month recall rate was good (higher than 80%), there was likely to be an element of selection bias in that paired CAT scores were not obtained in those who had died or were receiving inpatient treatment at the recall timepoint. Nevertheless, it was reassuring to show that change in the CAT correlated significantly with change in SGRQ and physical performance measures over 12 months.

The determination of the MCID remains controversial with no firm consensus, but is important in the validation of clinical instruments and the assessment of clinical studies. Two main approaches are generally used: anchor and distribution-based methods. Anchor-based methods rely on comparison of the change in outcome of interest with another outcome measure of change, known as the anchor or external criterion. However, this comparison is only relevant if there is an established association between the outcome of interest and the anchor. No consensus exists regarding the threshold strength of the association: some investigators have suggested an arbitrary minimum correlation coefficient of greater than 0·50, although others have suggested 0·30.[11,20] Although cross-sectional studies have shown strong correlations of the CAT and SGRQ, with four units on the SGRQ corresponding to 1·6 on the CAT, this association might differ when assessing change in these instruments.[26] The MCID is most useful for clinicians or researchers in the "change" setting, and therefore we made great efforts to assess change in CAT in three independent cohorts, including longitudinal follow-up. The candidate anchors with strongest correlations were COPD-specific health status questionnaires (CRQ and SGRQ), presumably because these instruments measure similar constructs to the CAT, with correlation coefficients ranging from 0·32 to 0·47. Previous guidance has recommended the use of several approaches and triangulation of methods.[20] We used several anchors, adopted different methods of anchor-based estimations (linear regression analysis, sensitivity and specificity analysis using receiver operating characteristic curves), and presented change in CAT in three different clinical scenarios to provide clinical context.

The estimates of the MCID of the CAT were consistent across different cohorts in different scenarios over different timeframes (table 5), providing some degree of corroboration and credibility, although the correlations with the external anchors were only moderate and caution is needed in the interpretation. At the individual level (the CAT score permits only integers), the ROC analysis and the anchor responses to the GRCQ estimated the MCID to be –2. From the linear regression analysis, the population-level MCID estimates ranged from –1·2 to –2·8, with the mean of these estimates being –2·0. Therefore, both the population-level estimate (used to assess group effects of treatments or differences between populations) and the individual patient-level estimate (used for responder analysis) was –2.

Anchor-based methods to estimate MCID are often preferred to distribution-based approaches because they take into account patient-reported benefit or deterioration. However, there are limitations. As mentioned previously, anchor-derived estimates are only valid if the outcome of interest correlates with the anchor. Any patient-reported outcome recorded before and after a period of time is also subject to recall bias. Anchors are designed to detect change in outcome but rarely take into account costs to the

***Panel:* Research in context**

Systematic review

We searched PubMed for studies in English focusing on estimates of the minimum clinically important difference (MCID) of the COPD Assessment Test (CAT) from inception, to Oct 21, 2013. We used the terms "MID" or "minimal important difference" or "minimum important difference" or "MCID" or "minimal clinically important difference" and "CAT" or "COPD Assessment Test" or "Chronic Obstructive Pulmonary Disease Assessment Test". We identified three studies[7,24,25] that provided estimates of the MCID of the CAT, although through retrospective opportunistic post-hoc analyses. These studies looked at improvement in stable populations using only one methodological approach for estimation; the resulting estimates ranged from a –1·3 to –3·76 point change.

Interpretation

As far as we are aware, this is the first report to provide prospective data to estimate the MCID of the CAT. Our studies provide 19 separate estimates of the MCID using both distribution-based and anchor-based methods, from three separate clinically relevant cohorts, including stable patients having pulmonary rehabilitation and longitudinal follow-up, and also in those recovering from exacerbation. To our knowledge, Study 1 is the largest study using the CAT in pulmonary rehabilitation. Our findings suggest that a decrease in CAT of 2 points is the most reliable estimate of the MCID at the individual and population level. This estimate will allow clinicians to interpret clinically important change in individual patients, and help researchers with sample size calculations and the interpretation of CAT data in response to intervention studies.

patient, for example side-effects of therapy. Another limitation is that changes in an outcome measure might be associated with baseline level. Potential reasons for this association include so-called floor and ceiling effects. The CAT score ranges from 0 to 40 with a decrease in score showing improvement in health status. At extreme values (eg, a baseline CAT of 0), the instrument cannot detect an improvement in health status. Similarly, if the baseline CAT is 40, there is no room for further worsening in health status. Although we used three independent cohorts of patients with COPD, with varying baseline mean values, these were generally recruited from a secondary care setting. The present studies cannot address whether our estimates of MCID for the CAT hold true in populations at extremes of the health spectrum (for example, community-dwelling asymptomatic patients with COPD, or highly symptomatic patients receiving palliative care).

To provide a comprehensive approach, we also used distribution-based methods to estimate the MCID of the CAT. Distribution-based approaches use statistical methods, and are based on the distribution of the cohort and the reliability of the measure; as such they do not take into account whether the recorded change is important from the patient's perspective. Previous investigators have noted that 0·5 SD and the SEM might approximate the MCID in some patient-reported outcome measures,[21,22] although other measures have also been described including 1·96 SEM and minimum detectable change; depending on which distribution-based approach is used, a range of different estimates of the MCID can be generated which might significantly limit their interpretation.[27] Another specific limitation to the present studies was that the test-retest reliability of the CAT was assumed on the basis of previous studies of much shorter duration. Our estimates of the MCID for the CAT, using 0·5 SD and SEM, were consistent across the three clinical scenarios, suggesting that the distributions of our three cohorts were much the same. However, it was also noticeable that the distribution-based methods consistently estimated the MCID to be greater than the anchor-based estimates, suggesting that the CAT had a wide distribution in our cohorts. This situation is not unique in patients with COPD. For example, Puhan and colleagues[28] showed that the MCID for maximal cycle exercise capacity ranged from 2·2 W to 3·3 W, whereas distribution-based estimates were 5·3–5·5 W.[28]

Although we provide data from both approaches we have chosen to place greater emphasis on the anchor-based estimates for two reasons. First, although distribution-based estimates provide supportive information regarding a significant change, they do not provide a direct measure of minimum clinical importance, a view shared by other researchers.[28,29] Second, the MCIDs estimated by distribution-based methods were greater than the mean change identified after pulmonary rehabilitation, widely accepted as a highly effective intervention that significantly improves health-related quality of life in COPD, and the mean change identified during recovery after admission to hospital for severe exacerbation of COPD. In view of these clinical contexts, we believe the data derived from distribution-based methods provide information about clinical significance but might overestimate the true MCID.

In the present studies, our focus was on establishing the minimum clinically important improvement. In studies 1 and 2, the number of patients with improving health status far outweighed those who had worsening health status (90% *vs* 3% according to the GRCQ), and we did not report estimates of minimum important clinical deterioration for fear of imprecision. We believe that the magnitude of the minimum clinically important deterioration is likely to be much the same as for improvement. For example, when we compared the linear regression slopes of those who improved and deteriorated according to their anchor, we identified no significant difference. In Study 3, in which there were more equivalent numbers of patients improving and worsening their health status, the receiver operating characteristic curves estimated similar magnitude cutoffs. Further studies focusing on worsening health status are needed to confirm whether patients perceive size of deterioration differently to size of improvement.

The main aim of the present studies was to estimate the MCID in change in CAT, and hence paired measurements were used for the analyses. In all three studies, there were missing data (usually because of exacerbation) and it could be argued that these missing data might bias our estimates of the MCID. If a patient deteriorated because of exacerbation or admission to hospital, we would expect not only the CAT to worsen but also the external anchors. Because there was no difference in the slopes of change in patients who improved compared with those who deteriorated, we do not believe that missing data were a significant source of bias in our estimates of the MCID for the CAT.

In summary, the present studies show that the CAT is responsive to the effects of pulmonary rehabilitation and recovery from admission to hospital for acute exacerbation of COPD. By use of various health status and global rating of change questionnaires as external anchors, we estimate the minimum important improvement of the CAT to be a two point decrease at both the individual and population level. This information could be useful in the clinical interpretation of CAT data, particularly in response to intervention studies.

Contributors

All authors contributed to the analysis and interpretation of data, and the preparation of the report. WD-CM conceived the idea and is the guarantor of the paper, taking responsibility for the integrity of the work as a whole, from inception to published Article.

Conflicts of interest

MIP has received personal reimbursement for lecturing or consultancy regarding muscle function in COPD from Novartis and Philips Respironics; he discloses institutional reimbursement for

consultancy from GlaxoSmithKline, Novartis, Regeneron, Lilly, Biomarin, and Boehringer Ingelheim and institutional agreements to do research with GlaxoSmithKline, Novartis, AstraZeneca, and Philips Respironics. All other authors declare that they have no conflicts of interest.

Acknowledgments

SSCK is supported by the Medical Research Council. WD-CM is supported by a National Institute for Health Research Clinician Scientist Award, a Medical Research Council (UK) New Investigator Research Grant, and a National Institute for Health Research Clinical Trials Fellowship. This project was done at the NIHR Respiratory Biomedical Research Unit at the Royal Brompton and Harefield NHS Foundation Trust and Imperial College London; the salaries of JLC, SEJ, and MIP are wholly or partly funded by the NIHR Biomedical Research Unit. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research, nor the Department of Health. We thank the Harefield Pulmonary Rehabilitation Team and the Hillingdon COPD Outreach Team for their help in recruiting participants.

References

1 Ringbaek T, Martinez G, Lange P. A comparison of the assessment of quality of life with CAT, CCQ, and SGRQ in COPD patients participating in pulmonary rehabilitation. *COPD* 2012; **9:** 12–15.

2 Jones PW, Harding G, Berry P, Wiklund I, Chen WH, Kline Leidy N. Development and first validation of the COPD Assessment Test. *Eur Respir J* 2009; **34:** 648–54.

3 Jones PW, Harding G, Wiklund I, et al. Tests of the responsiveness of the COPD assessment test following acute exacerbation and pulmonary rehabilitation. *Chest* 2012; **142:** 134–40.

4 Agusti 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:** 492–98.

5 Miravitlles M, Garcia-Sidro P, Fernandez-Nistal A, Buendia MJ, Espinosa de Los Monteros MJ, Molina J. Course of COPD assessment test (CAT) and clinical COPD questionnaire (CCQ) scores during recovery from exacerbations of chronic obstructive pulmonary disease. *Health Qual Life Outcomes* 2013; **11:** 147.

6 Dodd JW, Marns PL, Clark AL, et al. The COPD Assessment Test (CAT): short- and medium-term response to pulmonary rehabilitation. *COPD* 2012; **9:** 390–94.

7 Dodd JW, Hogg L, Nolan J, et al. The COPD assessment test (CAT): response to pulmonary rehabilitation. A multicentre, prospective study. *Thorax* 2011; **66:** 425–29.

8 Mackay AJ, Donaldson GC, Patel AR, Jones PW, Hurst JR, Wedzicha JA. Usefulness of the Chronic Obstructive Pulmonary Disease Assessment Test to evaluate severity of COPD exacerbations. *Am J Respir Crit Care Med* 2012; **185:** 1218–24.

9 Vestbo J, Hurd SS, Agusti AG, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2013; **187:** 347–65.

10 Schunemann HJ, Puhan M, Goldstein R, Jaeschke R, Guyatt GH. Measurement properties and interpretability of the Chronic respiratory disease questionnaire (CRQ). *COPD* 2005; **2:** 81–89.

11 Schunemann HJ, Griffith L, Jaeschke R, Goldstein R, Stubbing D, Guyatt GH. Evaluation of the minimal important difference for the feeling thermometer and the St. George's Respiratory Questionnaire in patients with chronic airflow obstruction. *J Clin Epidemiol* 2003; **56:** 1170–76.

12 Jones PW. St George's respiratory questionnaire: MCID. *COPD* 2005; **2:** 75–79.

13 Pauwels RA, Buist AS, Calverley PM, Jenkins CR, Hurd SS. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. *Am J Respir Crit Care Med* 2001; **163:** 1256–76.

14 Williams JE, Singh SJ, Sewell L, Guyatt GH, Morgan MD. Development of a self-reported Chronic Respiratory Questionnaire (CRQ-SR). *Thorax* 2001; **56:** 954–59.

15 Kon SS, Patel MS, Canavan JL, et al. Reliability and validity of 4-metre gait speed in COPD. *Eur Respir J* 2013; **42:** 333–40.

16 Jones SE, Kon SS, Canavan JL, et al. The five-repetition sit-to-stand test as a functional outcome measure in COPD. *Thorax* 2013; **68:** 1015–20.

17 Singh SJ, Morgan MD, Scott S, Walters D, Hardman AE. Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax* 1992; **47:** 1019–24.

18 Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific Quality of Life Questionnaire. *J Clin Epidemiol* 1994; **47:** 81–87.

19 Kon SS, Clark AL, Dilaver D, et al. Response of the COPD Assessment Test to pulmonary rehabilitation in unselected chronic respiratory disease. *Respirology* 2013; **18:** 974–77.

20 Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008; **61:** 102–09.

21 Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003; **41:** 582–92.

22 Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol* 1999; **52:** 861–73.

23 Bolton CE, Bevan-Smith EF, Blakey JD, et al. British Thoracic Society guideline on pulmonary rehabilitation in adults. *Thorax* 2013; **68** (suppl 2)**:** ii1–30.

24 Tsiligianni IG, van der Molen T, Moraitaki D, et al. Assessing health status in COPD. A head-to-head comparison between the COPD assessment test (CAT) and the clinical COPD questionnaire (CCQ). *BMC Pulm Med* 2012; **12:** 20.

25 Jones PW, 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.

26 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.

27 Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 2007; **7:** 541–46.

28 Puhan MA, Chandra D, Mosenifar Z, et al. The minimal important difference of exercise tests in severe COPD. *Eur Respir J* 2011; **37:** 784–90.

29 Rennard SI. Minimal clinically important difference, clinical perspective: an opinion. *COPD* 2005; **2:** 51–55.