Qual Life Res (2015) 24:2345–2354
DOI 10.1007/s11136-015-0983-x

Does the COPD assessment test (CAT™) questionnaire produce similar results when self- or interviewer administered?

A. Agusti[1,2] · J. J. Soler-Cataluña[3] · J. Molina[4] · E. Morejon[5] · M. Garcia-Losa[5,6] · M. Roset[6] · X. Badia[6]

Accepted: 1 April 2015 / Published online: 7 April 2015

Abstract
Purpose The COPD assessment test (CAT) is a questionnaire that assesses the impact of chronic obstructive pulmonary disease (COPD) on health status, but some patients have difficulties filling it up by themselves. We examined whether the mode of administration of the Spanish version of CAT (self vs. interviewer) influences its scores and/or psychometric properties.
Methods Observational, prospective study in 49 Spanish centers that includes clinically stable COPD patients ($n = 153$) and patients hospitalized because of an exacerbation (ECOPD; $n = 224$). The CAT was self-administered (CAT-SA) or administered by an interviewer (CAT-IA) based on the investigator judgment of the patient's capacity. To assess convergent validity, the Saint George's Respiratory Disease Questionnaire (SGRQ) and the London Chest Activity of Daily Living (LCADL) instrument were also administered. Psychometric properties were compared across modes of administration.
Results A total of 118 patients (31 %) completed the CAT-SA and 259 (69 %) CAT-IA. Multiple regression analysis showed that mode of administration did not affect CAT scores. The CAT showed excellent psychometric properties in both modes of administration. Internal consistency coefficients (Cronbach's alpha) were high (0.86 for CAT-SA and 0.85 for CAT-IA) as was test–retest reliability (intraclass correlation coefficients of 0.83 for CAT-SA and CAT-IA). Correlations with SGRQ and LCADL were moderate to strong both in CAT-SA and CAT-IA, indicating good convergent validity. Similar results were observed when testing longitudinal validity.
Conclusions The mode of administration does not influence CAT scores or its psychometric properties. Hence, both modes of administration can be used in clinical practice depending on the physician judgment of patient's capacity.

Keywords COPD · Quality of life · Questionnaire · CAT · Mode of administration

Introduction

The new recommendations of the Global Initiative for Chronic Obstructive Lung Disease (GOLD) emphasize the importance of symptom assessment in the management and treatment of patients with chronic obstructive pulmonary disease (COPD) [1]. To this end, GOLD recommends the use of the modified Medical Research Council (mMRC) breathlessness scale, the COPD assessment test (CAT), and/or the Clinical COPD Questionnaire (CCQ) (www.goldcopd.org). The CAT™ was originally developed as a self-administered questionnaire that includes eight items that assess symptoms, limitation of daily activity, sleep quality, and energy and that provides a single score [2, 3]. The CAT logo is a trade mark of the GlaxoSmithKline group of companies; the CAT can be freely used. The CAT

✉ M. Garcia-Losa
mgarcialosa@es.imshealth.com

1 Thorax Institute, Hospital Clinic, IDIBAPS, University of Barcelona, Barcelona, Spain

2 FISIB, CIBERES, Mallorca, Spain

3 Hospital de Requena, Requena, Valencia, Spain

4 Centro de salud Francia, Fuenlabrada, Madrid, Spain

5 GlaxoSmithKline, Tres Cantos, Madrid, Spain

6 IMS Health, C/Dr. Ferran 25-27, 08034 Barcelona, Spain

has been extensively tested [4–8] and has been shown to be reliable, valid, and sensitive to change during exacerbations of the disease (ECOPD) [9, 10]. In clinical practice, however, some patients with poor health status and/or a low educational level may be unable to complete the questionnaire by themselves and need support from an interviewer. Whether CAT scores and/or its psychometric properties change if it is self-completed or if administered by healthcare personnel has not been investigated before. The answer to this question is not straightforward since some [11–14] but not all [15–18] previous studies investigating this same issue in different questionnaires reported a significant influence of the mode of administration on scores. Likewise, it is also relevant to investigate whether the mode of administration may affect the psychometric properties of CAT [14, 19], particularly its reliability (i.e., internal consistency and/or test–retest reliability) and validity, the latter including convergent validity (the extent to which the instrument in question correlates in a consistent fashion with instruments measuring similar constructs, as well as with more clinical measures), known groups' validity (the capacity to discriminate between patients classified according to clinically relevant variables), and longitudinal validity (i.e., the correlation in changes over time observed on different instruments purporting to measure similar constructs) [20].

In this study, we hypothesized that the questionnaire's mode of administration, self-administered (SA) versus interviewer administered (IA), would not significantly affect CAT scores or the instrument's psychometric properties in COPD patients, both clinically stable and hospitalized because of an exacerbation of the disease (ECOPD).

Methods

Study design and ethics

Results presented here correspond to a post hoc analysis of the validation study of the Spanish version of the CAT published before [9]. This was an observational, prospective, and longitudinal study in which 49 centers from around Spain participated. They included both pulmonary services and primary care surgeries; all of them well experienced in the treatment of COPD (Appendix). In brief, two groups of patients were recruited consecutively between December 2009 and April 2010. The first one included patients hospitalized because of ECOPD who were visited twice during the study, once during the first 48 h of hospitalization and then 4 ± 1 weeks after hospital discharge; the second one included clinically stable patients, who were also visited twice, at recruitment and 4 ± 1 weeks later [9].

The study protocol was approved by the Clinical Research Ethics Committee of the Hospital Clinic in Barcelona (Spain), and all patients signed informed consent to participate.

Participants

Patients were included in the study if they: (1) were older than 40 years of age, (2) were current or former smokers (>10 pack-years), and (3) suffered COPD according to the diagnostic criteria of the GOLD recommendations (FEV1/FVC post-bronchodilator <0.7) [1]. Patients were considered clinically stable if they were free of ECOPD and/or any change in treatment during the 8 weeks prior to inclusion. Patients were excluded if they suffered from any other uncontrolled or severe concomitant illness. Sample size was calculated as reported previously [9].

Measurements

The Spanish translation of the CAT questionnaire used in the present study was developed through a standard forward and back-translation process and, like the original English version, includes eight items (severity of cough, presence of mucus, chest tightness, dyspnea, limitations during domestic activities, social limitations, sleep, and energy restriction), each with a response scale ranging from 0 (no limitations) to 5 (very limited) [9]. The overall score also ranges from 0 (no effect on health status) to 40 (considerable effect on health status). In clinical practice, clinicians decide whether to ask the patient to self-administer the CAT or to administer it by interview based on their judgment of individual patient capacity. To mimic this situation, in this study patients were assigned to either interview CAT administration (CAT-IA) or self-administration (CAT-SA) based on investigator opinion of the patient's capacity to self-complete. In the former case, the interviewer was a health professional (physician or nurse) who had been previously trained for this purpose.

To test the convergent validity of both the CAT-IA and CAT-SA, other HRQoL measures were a version of the Saint George's Respiratory Disease Questionnaire adapted for patients with COPD (SGRQ-C) [21] and the London Chest Activity of Daily Living (LCADL) scale [22]. Both have been validated in Spanish [23, 24].

As published before [9], other variables determined in the study included age, gender, level of education, smoking history (including number of years smoking, or since quitting), as well as total smoking exposure (pack-years), presence of concomitant illnesses, date of last ECOPD, number of ECOPD in the past 6 months, presence and number of ECOPD during the study period, and medication at baseline. Dyspnea was assessed by the MRC

breathlessness scale [25]. Forced spirometry was determined according to international guidelines [26] at recruitment for stable patients and during the second visit for ECOPD patients. If a technically valid forced spirometry had been obtained during the 3 months prior to entering the study, this was considered acceptable and spirometry was not repeated.

Statistical analysis

Results are presented as absolute numbers (proportions) or means (standard deviations), as appropriate. Baseline clinical and socio-demographic characteristics of the CAT-SA and CAT-IA groups were compared using parametric (Student's t test and ANOVA) and nonparametric (Mann–Whitney *U*, Kruskal–Wallis, Chi-squared) tests. Potential confounders on the effect of the mode of administration on CAT scores were explored using multiple regression analysis where the dependent variable was the CAT score at baseline and the independent ones were sex, age, mode of administration, level of education, overall clinical assessment, MRC breathlessness score, group (stable or ECOPD), and presence of exacerbations. To determine whether mode of administration affected the questionnaire's psychometric performance, the following attributes and psychometric properties of the CAT were examined separately for the CAT-SA and CAT-IA groups. Feasibility was tested by analyzing the proportion of patients with complete scores on the questionnaire. Ceiling and floor effects (i.e., the percentage of patients scoring 100 and 0, respectively) were calculated. Reliability was tested by determining internal consistency (using Cronbach's α) in all patients, and test–retest reliability in stable patients (using the intraclass correlation coefficient or ICC). For both Cronbach's α and the ICC, values over 0.7 were considered acceptable [27].

Known groups' validity was investigated by comparing CAT scores in COPD patients according to clinical stability, GOLD stages, MRC dyspnea scale categories, and presence and number of exacerbations in the previous 6 months [20]. Convergent validity was tested by examining correlations (using Pearson's correlation coefficient) between CAT-IA or CAT-SA scores and mean SGRQ-C and LCADL total and dimension scores at baseline, and by examining correlations with FEV1 readings and number of exacerbations in the previous 6 months. Correlation coefficients were compared according to administration mode using the Fisher r-to-z transformation test. We expected moderate to strong correlations between the CAT and the SGRQ-C and LCADL whichever mode of administration was used. Finally, we investigated whether the instrument's longitudinal validity (i.e., the degree to which changes in CAT scores over time correlated with changes in clinical and other outcome measures [14]) was affected by mode of administration by examining and comparing correlations between changes in CAT scores and changes in FEV1, SGRQ-C, LCADL, and MRC dyspnea scale scores in the two study groups (CA-SA and CAT-IA). To this end, only data from patients using the same mode of CAT administration in both study visits (91.3 %) were used. All analyses were performed using SPSS 15.0 for Windows. Statistical significance was set at 0.05.

Results

Patient characteristics

A total of 377 patients (153 with clinically stable COPD and 224 with ECOPD) were included in the analysis, of whom 118 (31 %) completed the questionnaire by self-administration and 259 (69 %) by interview administration at the recruitment visit. The CAT-SA group included 63 (53 %) stable patients and 55 (47 %) ECOPD patients, whereas the CAT-IA group included 90 (35 %) stable and 169 (66 %) ECOPD patients. A minority of patients ($n = 30$, 8.7 %) switched the mode of CAT administration at the follow-up visit (19 CAT-SA and 11 CAT-IA).

Table 1 presents the main socio-demographic, clinical, and functional characteristics of the CAT-SA and CAT-IA groups. Patients in the CAT-SA group were younger (67 vs. 72 years, $p < 0.01$), had a higher proportion of females (15.3 vs. 3.5 %, $p < 0.01$), and, as expected, had a higher educational level. They also reported less dyspnea (MRC), had better health status (SGRQ-C) (46 vs. 57, $p < 0.01$), and referred fewer ECOPD in the previous 6 months (39 vs. 62 %, $p < 0.01$). Only one patient (in the CAT-IA group) failed to complete all items of the questionnaire, indicating excellent feasibility.

CAT scores by mode of administration

Table 2 shows the score distribution according to administration mode and disease status (active vs. stable). Raw scores were higher in CAT-IA and in patient with active disease, reflecting poorer health status. The ceiling and floor effects (i.e., the % of patients with the maximum and minimum possible scores, respectively) were exceptional, with negligible or nonexistent floor and ceiling effects in both groups. However, after adjusting for differences in socio-demographic and clinical variables between the CAT-IA and CAT-SA groups in the multiple regression analysis (Table 3), the effect was no longer statistically significant (β −1.696; $p = 0.059$). The independent variable showing the greatest impact on CAT scores was

Table 1 Socio-demographic, clinical, and functional characteristics in the overall population and by mode of CAT administration

	Overall ($n = 377$)	CAT-SA ($n = 118$)	CAT-IA ($n = 258$)	p value
Age, mean (SD), years	69.9 (9.4)	66.8 (9.8)	71.3 (8.9)	<0.01
Male, *n* (%)	350 (92.8 %)	100 (84.7 %)	250 (96.5 %)	<0.01
Highest educational level, *n* (%)*				<0.01
No formal education	98 (26.2 %)	12 (10.3 %)	86 (33.5 %)	
Primary	207 (55.3 %)	64 (54.7 %)	143 (55.6 %)	
Secondary	36 (9.6 %)	18 (15.4 %)	18 (7.0 %)	
University	33 (8.8 %)	23 (19.7 %)	10 (3.9 %)	
Comorbidities, *n* (%)				NS
Any	311 (82.5 %)	101 (85.6 %)	210 (81.1 %)	
Type 1 diabetes mellitus	8 (2.1 %)	3 (2.5 %)	5 (1.9 %)	
Type 2 diabetes mellitus	89 (23.6 %)	19 (16.1 %)	70 (27.0 %)	
Ischemic heart disease	45 (11.9 %)	12 (10.2 %)	33 (12.7 %)	
Hypertension	176 (46.7 %)	47 (39.8 %)	129 (49.8 %)	
Other	237 (62.9 %)	83 (70.3 %)	154 (59.5 %)	
Years since diagnosis COPD, mean (SD)	9.6 (7.4)	9.3 (9.3)	9.7 (6.3)	NS
Study group, *n* (%)				<0.01
Stable	153 (40.6 %)	63 (53.4 %)	90 (34.7 %)	
Acute	224 (59.4 %)	55 (46.6 %)	169 (65.3 %)	
Patients reporting exacerbations in the 6 months prior to study entry, *n* (%)	206 (54.6 %)	46 (39.0 %)	160 (61.8 %)	<0.01
Number of exacerbations in the 6 months prior to study entry, mean (SD)	2.4 (2.0)	2.0 (1.6)	2.5 (2.1)	NS
MRC dyspnea scale, *n* (%)				<0.01
Grade 0	19 (5.1 %)	13 (11.0 %)	6 (2.3 %)	
Grade I	107 (28.5 %)	43 (36.4 %)	64 (24.8 %)	
Grade II	111 (29.5 %)	34 (28.8 %)	77 (29.8 %)	
Grade III	81 (21.5 %)	18 (15.3 %)	63 (24.4 %)	
Grade IV	58 (15.4 %)	10 (8.5 %)	48 (18.6 %)	
SGRQ-C questionnaire, overall mean (SD) score	53.8 (23.7)	46.0 (21.9)	57.2 (23.7)	<0.01
LCADL questionnaire, overall mean (SD) score	26.8 (16.1)	24.5 (15.5)	27.9 (16.3)	NS

NS not significant, *NA* not Applicable, *SD* standard deviation

* Missing data from 3

degree of dyspnea as measured with the MRC breathlessness scale.

Figure 1 shows the mean CAT scores by mode of administration in several clinically meaningful groups (clinical stability vs. ECOPD, severity of airflow limitation according to the GOLD grades, degree of dyspnea by MRC scores, and presence/absence of previous ECOPD in the previous 6 months). CAT-IA and CAT-SA mean scores were significantly different when patients were grouped by MRC dyspnea score and by the presence of previous exacerbations. It is also worth noting that CAT-IA scores were always numerically higher than CAT-SA ones in all comparisons (Fig. 1).

Influence of mode of administration on CAT psychometric properties

Internal consistency for the CAT-SA and CAT-IA is shown in Table 2 according to patient type (stable or ECOPD). Internal consistency for the questionnaire was good whichever mode of administration was used, with Cronbach's alphas of 0.86 and 0.85 for the CAT-SA and CAT-IA groups, respectively. Test–retest reliability was also satisfactory for the instrument, with an ICC of 0.83 in patients with stable COPD for both the CAT-SA and the CAT-IA.

In terms of convergent validity (see Table 4), correlations between CAT-SA and CAT-IA scores and SGRQ and

Table 2 Description of CAT scores internal consistency by mode of administration

	CAT baseline score*							
	Valid *N*	Mean (SD)	Min	Median	Max	Floor effect	Ceiling effect	Internal consistency
Mode of administration								
SA								
Stable	$N = 63$	13.2 (6.9)	0.0	12.0	28.0	0.8 %	–	0.86
Acute	$N = 55$	20.7 (7.6)	4.0	20.0	37.0			
IA								
Stable	$N = 90$	17.7 (8.4)	1.0	16.0	36.0	–	0.8 %	0.85
Acute	$N = 168$	23.0 (8.5)	5.0	23.0	40.0			
Total	$N = 376$	19.7 (8.9)	0.0	19.0	40.00	0.3 %	0.5 %	0.86

SA self-administered, *IA* interviewer administered

* 0: no impact on HRQL; 40: highest impact on HRQL

Table 3 Linear regression model results (CAT score = dependent variable)

	Parameter	*B*	SE	*t*	Sig.	95 % confidence interval	
						Lower limit	Upper limit
	Intersection	25.266	3.411	7.406	.000	18.555	31.978
Sex	Male	−1.204	1.509	−.798	.426	−4.173	1.765
	Female	0[a]	–	–	–	–	–
Age		−.019	.042	−.449	.654	−.101	.064
Mode of administration	Self	−1.696	.895	−1.896	.059	−3.456	.064
	Interview	0[a]	–	–	–	–	–
MRC dyspnea scale grades	0	−13.149	2.158	−6.093	.000	−17.394	−8.903
	1	−9.181	1.409	−6.514	.000	−11.953	−6.408
	2	−5.213	1.277	−4.083	.000	−7.724	−2.701
	3	−1.131	1.306	−.866	.387	−3.701	1.439
	4	0[a]	–	–	–	–	–
Group	Acute	2.611	.922	2.832	.005	.797	4.424
	Stable	0[a]	–	–	–	–	–
Highest education level	No formal education	3.211	1.612	1.992	.047	.040	6.382
	Primary	.911	1.448	.629	.530	−1.938	3.760
	Secondary	1.581	1.771	.892	.373	−1.904	5.066
	University	0[a]	–	–	–	–	–
Presence of exacerbations	No	−1.737	.835	−2.079	.038	−3.380	−.094
	Yes	0[a]	–	–	–	–	–
GOLD grades	1	−1.846	1.556	−1.186	.236	−4.907	1.216
	2	−2.357	1.229	−1.917	.056	−4.775	.062
	3	−1.348	1.075	−1.254	.211	−3.463	.767
	4	0[a]	–	–	–	–	–

[a] Reference level

LCADL overall scores were similar, without statistically significant differences, and generally strong (0.81 with the SGRQ overall score for both the CAT-SA and CAT-IA and 0.61 and 0.65 with the LCADL overall score for the CAT-SA and CAT-IA, respectively). All correlations were also significant at $p < 0.01$. Neither the CAT-IA nor CAT- A showed a clear advantage in terms of correlations with individual dimensions of the SGRQ and LCADL with correlations for both instruments being generally moderate to strong. Only the CAT-IA showed statistically significant correlations with clinical measures of COPD ($r = -0.28$ with pre-bronchodilator %FEV1 and $r = -0.39$ with post-bronchodilator %FEV1 and %FEV1/FCV; $p \leq 0.01$). CAT-SA and CAT-IA scores generally showed weak to

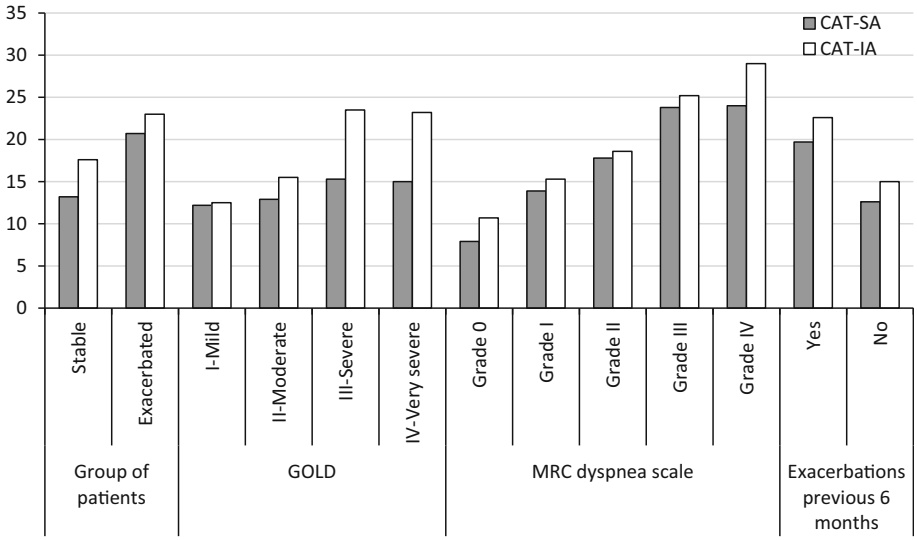

Fig. 1 Mean CAT score by mode of administration in clinically meaningful groups: stable versus ECOPD, severity of airflow limitation (GOLD grades), degree of dyspnea (MRC), and presence/absence of previous ECOPD. For further explanations, see text

moderate correlations with the number of ECOPD episodes during the 6 months prior to inclusion in the study (Pearson's $r = 0.44$ for CAT-SA and 0.37 for CAT-IA, with all correlations significant at $p < 0.05$).

Similar results were observed in the analysis of longitudinal validity, whereby changes in CAT-SA and CAT-IA scores between the two study visits correlated moderately or strongly with changes in SGRQ and LCADL overall and dimension scores (e.g., r of 0.67 and 0.47 for correlation between changes in SGRQ overall score and CAT-IA and CAT-SA scores, respectively, and r of 0.43 and 0.45 for correlation between changes in LCADL overall score and CAT-IA and CAT-SA scores). Correlations with FEV1 were again much weaker and only statistically significant in the case of the CAT-IA (r of 0.25 and 0.19 with pre-bronchodilator %FEV1 and post-bronchodilator %FEV1, respectively; $p < 0.05$).

Discussion

This study shows that the way of administration of CAT (self- vs. aided administration), as decided by the attending clinician, does not influence significantly the scores and psychometric properties of CAT.

Previous studies

Many previous studies have investigated different methodological and/or clinical aspects of CAT [4–10]. To our knowledge, however, this is the first one to explore whether its mode of administration (self-administered vs. interviewer administered) affects its scores and/or psychometric properties. This may be clinically relevant since a recent study reported that more than 50 % of COPD patients required assistance to self-complete the CAT [28]. Further, it is important to investigate potential effects of a particular mode of administration on scoring, to ascertain whether it is valid to compare scores or pool data obtained using different modes of administration [29]. For instance, previous studies reported that the mode of administration can affect scores on the emotional dimensions of health-related quality of life (HRQOL) scales but not the physical dimensions [11, 30] and that interview administration leads to scores reflecting better health [13, 14], whereas others have found a negligible or nonexistent effects [15–18].

Interpretation of results

To mimic real life conditions, patients were not randomized to SA or IA groups. Rather, they were assigned to one of these two groups based on the judgment of each clinical investigator. As a result, the baseline characteristics of both groups were not entirely comparable (Table 1). As expected, patients in the SA group had a higher level of education and better health status, and they reported fewer exacerbations in the 6 months before being included in the study.

Examination of raw CAT scores (Fig. 1) suggested that interview administration might lead to slightly higher ones (i.e., poorer health status). After adjusting for differences in baseline characteristics (Table 1) of the two groups (self- and interviewer administered groups), differences were almost statistically significant, but they were no longer so (Table 3). This may be relevant in COPD patients, since their ability to self-administer the questionnaire may be affected by the severity of the disease or by the low educational level of some patients [28]. The fact that the

Table 4 Convergent validity (correlation between changes in CAT and SGRQ-C and LCADL scores) by mode of administration

Questionnaire and dimensions for correlation	CAT overall	CAT-SA	CAT-IA	*p* value (two-sided test)
SGRQ-C—symptoms				
r	0.65	0.57	0.68	0.11
p	<0.01	<0.01	<0.01	
n	376	118	258	
SGRQ-C—activities				
r	0.73	0.76	0.71	0.35
p	<0.01	<0.01	<0.01	
n	360	110	250	
SGRQ-C—impact				
r	0.80	0.76	0.80	0.37
p	<0.01	<0.01	<0.01	
n	365	111	254	
SGRQ-C—total				
r	0.82	0.81	0.81	1.00
p	<0.01	<0.01	<0.01	
n	359	110	249	
LCADL—self-care				
r	0.63	0.54	0.64	0.18
p	<0.01	<0.01	<0.01	
n	365	113	252	
LCADL—domestic activities				
r	0.40	0.54	0.37	0.06
p	<0.01	<0.01	<0.01	
n	350	110	240	
LCADL—physical activities				
r	0.52	0.58	0.49	0.27
p	<0.01	<0.01	<0.01	
n	368	114	254	
LCADL—leisure activities				
r	0.59	0.49	0.62	0.10
p	<0.01	<0.01	<0.01	
n	369	115	254	
LCADL—total				
r	0.63	0.65	0.61	0.58
p	<0.01	<0.01	<0.01	
n	344	107	237	

r Pearson's correlation coefficient, *p* *p* value, *n* number of valid cases

CAT is robust to the mode of administration also implies that data obtained using the self- or interview-administered versions can be pooled and compared. Although we did not test other possible administration effects, such as the impact of administering the instrument by telephone or computer, the recent review by Hood et al. [29] suggests that these were of considerably less importance than self- versus interview administered.

The mode of administration also had little effect on the questionnaire's psychometric performance, with the CAT showing highly satisfactory levels of reliability and validity in both forms of administration. Feasibility, floor and ceiling effects, internal consistency, and test–retest reliability were all well above accepted thresholds [27], while tests of convergent validity showed that the instrument correlated consistently with measures evaluating similar constructs (SGRQ-C and LCADL) whichever mode of administration was used. Equally positive results were seen for known groups' and longitudinal validity. The findings were very similar to those observed in a

comparison of the IA and SA versions of the Chronic Respiratory Questionnaire [14]. In the present study, the only minor difference detected between the two forms of administration was that the CAT-IA, but not the CAT-SA, showed statistically significant correlations with clinical variables used to measure lung function, such as pre- and post-bronchodilator %FEV1 and %FEV1/FCV. This may reflect an interviewer bias if healthcare professionals were aware of patients' lung function.

Potential limitations

Our study has several potential limitations that deserve discussion. First, patients were not randomized to the two different modes of administration. Instead, the method of administration was based on clinician judgment regarding the patient's ability to self-administer the instrument. This was done on purpose to mimic as much as possible a real life scenario. To address this bias, multivariate regression analysis was used to adjust for baseline differences between groups. However, we cannot exclude the possibility that unobserved factors (thus, not taken into account in the analysis) may have nevertheless influenced the results. Of note, though, Hood et al. [29] found that the fact that participants were randomized or not to a particular mode of questionnaire administration did not have a significant effect on survey responses. Second, the study sample size [9] was not calculated to assess the CAT performance using the two different modes of administration. Therefore, results should be confirmed in further studies.

Another potential limitation has to do with the approximate 9:1 male/female ratio found in our population which is indeed different from recent epidemiological data in Spain, where the EPI-SCAN study [31] found an approximate ratio of 3:1 among general population. Considering that our study subjects were recruited among consecutive COPD patients visiting the investigators (or being hospitalized), we can only hypothesize that the growing rate of female COPD does not produce, yet, rates of severe COPD as high as in male subjects. Nevertheless, although there does not seem to be any reason to expect a different behavior of males versus females with respect to the purpose of this study, especially considering that educational level is similar across both genders in Spain, given the small sample of female patients, this fact has to be taken into consideration when comparing psychometric differences between self-administered and interviewer administered.

Finally, due to the small number of patients who used a different mode of administration in the two study visits, we were not able to determine whether this switching affected scores and reliability. Until such analysis can be done, it would be advisable to ensure that the same mode of administration is used in different visits in any longitudinal study, especially because previous research indicated that reliability is higher when the same mode of administration is used at two recurrent visits [13].

Conclusions

CAT scores are not significantly influenced by the method of administration (self vs. aided), as decided by the attending clinician, and the instrument's psychometric properties were highly satisfactory whichever mode of administration was used. The results suggest that the way CAT is administered can be adapted to account for linguistic, cultural, educational, and functional diversity of patient populations without influencing scores or compromising the instrument's psychometric performance [15]. Further, they indicate that data obtained using different modes of administration can be pooled or compared and that either mode of administration is likely to give robust, reliable results. However, at this time it should still be recommended that patients self-complete the CAT whenever possible, leaving interviewer-led administration to cases when patients are unable to do so. Further research is required to confirm these observations and to determine whether sensitivity to change is affected by mode of administration.

Acknowledgments Authors thank all investigators who participated in the data collection of the study.

Conflict of interest Alvar Agustí has received from GSK (the company that has developed CAT) honorarium for speaking at meetings and participation on advisory boards on the CAT questionnaire. Juan J. Soler-Cataluña has received from GSK (the company that has developed CAT) honorarium for speaking at meetings and advisory board. Jesús Molina has received from GSK (the company that has developed CAT) honorarium for speaking and participation on advisory boards on the CAT questionnaire. Elena Morejón is currently employed by GlaxoSmithKline, she has been employed by this company for the last 5 years, and she holds stocks and share options in GlaxoSmithKline. Manuel García-Losa was an employee at GlaxoSmithKline at the time of this study conduction and owns GSK shares. Montserrat Roset and Xavier Badia have no conflict of interest to declare. The study was funded by GSK. The authors are responsible for the content and writing of the paper.

Appendix: List of participating pulmonary specialists and primary care physicians

Pulmonary specialists

Dr. Esteban Gonzalez, Hospital de Galdakao; Dr. Santos Perez, Hospital Universitario Bellvitge; Dr. Ausin Herrero, Hospital del Mar; Dr. Corral Penafiel, Hospital San Pedro Alcantara; Dr. Izquierdo Alonso, Hospital de Guadalajara; Dr. Laparra Galindez, Hospital Donostia; Dr. Calle Rubio,

Hospital Clinico San Carlos; Dr. Gonzalez Barcala, C.H. Pontevedra; Dr. Ferrer Sancho, H. Vall Hebron; Dr. Golpe Gomez, Hospital Xeral-Calde; Dr. Martinez Moragon, Hospital de Sagunto; Dr. Roig Figueroa, H. Universitario de Valladolid; Dr. Ancoechea Bermuda, Hospital de la Princesa; Dr. Garcia Rio, Hospital La Paz; Dr. Melero, H. Doce de Octubre; Dr. Almagro Mena, H. Mutua Terrassa; Dr. Carrizo Sierra, Hospital Miguel Servet; Dr. Jimenez Lozano; Hospital Baza; Dr. Martin Escudero, Hospital U. Rio Hortera; Dr. Peces-Barba, Fundacion Jimenez Diaz; Dr. Echave Sustaeta, Hospital Quiron- Madrid. Dr. Mengibar Vallejo, Hospital Baza; Dr. de Torres Tajes, Clinica Universidad Navarra; Dr. Ortega Ruiz, Hospital Virgen del Rocio; and Dr. Valdes Cuadrado, Hospital de Conxo.

Primary care physicians

Dr. Sanchez Gutierrez, C.S. Apolinar Moreno; Dr. Laporta Crespo, La Roda Centro de Salud; Dr. Medel Rocandio, San Miguel de Basauri; Dr. Molina Paris, C.S. Francia; Dr. Mate Sanchez, C.Salud Carballino; Dr. Lopez Rodriguez, C.S. Begonte; Dr. Lopez Peral, C.S. El Palo; Dr. Cimas Hernando, C.S. Contrueces; Dr. Canellas Isem, CAP Balsareny; Dr. Alonso Algorta, Bizcaia; Dr. Lopez Caro, C.S. Cotolino-Aguera; Dr. Nogueiras Santas, C.S. Val Minor; Dr. Quintano Jimenez, C.S.Lucena; Dr. Sellares Torres, Hospital Clinic (CAP Numancia); Dr. Carreras, CAP Sarria de Ter; Dr. Brau Tarrida, CAP La Mina; Dr. Martin Almendros, C.S. Burlada; Dr. Calvo Corbella, C.S. Pozuelo; Dr. Espigares Arroyo, CAP La Paz; Dr. Fernandez Barrial, C.S. de Blimea; Dr. Martinez Carrasco, Centro de Salud Fuencarral; Dr. Mora Moreno, C.S. El Molino de la Vega; Dr. Pascual Gil, C.S. Gusur; and Dr. Martin Perez, Cruce Arinaga.

References

1. Vestbo, J., Hurd, S. S., Agusti, A. G., Jones, P. W., Vogelmeier, C., Anzueto, A., et al. (2013). Global strategy for the diagnosis, management and prevention of Chronic Obstructive Pulmonary Disease, GOLD executive summary. *American Journal of Respiratory and Critical Care Medicine, 187*(4), 347–365.
2. Jones, P., Harding, G., Wiklund, I., Berry, P., & Leidy, N. (2009). Improving the process and outcome of care in COPD: development of a standardised assessment tool. *Prim Care Respir J, 18*, 208–215.
3. Jones, P. W., Harding, G., Berry, P., Wiklund, I., Chen, W.-H., & Kline Leidy, N. (2009). Development and first validation of the COPD assessment test. *European Respiratory Journal, 34*, 648–654.
4. Dodd, J. W., Hogg, L., Nolan, J., Jefford, H., Grant, A., Lord, V. M., et al. (2011). The COPD assessment test (CAT): Response to pulmonary rehabilitation. A multicentre, prospective study. *Thorax, 66*, 425–429.
5. Jones, P. W., Brusselle, G., Dal Negro, R. W., Ferrer, M., Kardos, P., Levy, M. L., et al. (2011). Properties of the COPD assessment test in a cross-sectional European study. *European Respiratory Journal, 38*, 29–35.
6. Jones, P. W., Tabberer, M., & Chen, W. H. (2011). Creating scenarios of the impact of COPD and their relationship to COPD assessment test (CAT^{TM}) scores. *BMC Pulmonary Medicine, 11*, 42.
7. Jones, P. W., Harding, G., Wiklund, I., Berry, P., Tabberer, M., Yu, R., & Leidy, N. K. (2012). Tests of the responsiveness of the Chronic Obstructive Pulmonary Disease (COPD) assessment test TM (CAT) following acute exacerbation and pulmonary rehabilitation. *Chest, 142*, 134–140.
8. Kelly, J. L., Bamsey, O., Smith, C., Lord, V. M., Shrikrishna, D., Jones, P. W., et al. (2012). Health status assessment in routine clinical practice: The chronic obstructive pulmonary disease assessment test score in outpatients. *Respiration, 84*, 193–199.
9. Agusti, A., Soler, J. J., Molina, J., Munoz, M. J., Garcia-Losa, M., Roset, M., et al. (2012). Is the CAT questionnaire sensitive to changes in health status in patients with severe COPD exacerbations? *Chronic Obstructive Pulmonary Disease, 9*, 1–7.
10. Mackay, A. J., Donaldson, G. C., Patel, A. R., Jones, P. W., Hurst, J. R., & Wedzicha, J. A. (2012). Usefulness of the chronic obstructive pulmonary disease assessment test to evaluate severity of COPD exacerbations. *American Journal of Respiratory and Critical Care Medicine, 185*, 1218–1224.
11. Cheung, Y. B., Goh, C., Thumboo, J., Khoo, K. S., & Wee, J. (2006). Quality of life scores differed according to mode of administration in a review of three major oncology questionnaires. *Journal of Clinical Epidemiology, 59*, 185–191.
12. Rhodes, T., Girman, C. J., Jacobsen, S. J., Guess, H. A., Hanson, K. A., Oesterling, J. E., & Lieber, M. M. (1995). Does the mode of questionnaire administration affect the reporting of urinary symptoms? *Urology, 46*, 341–345.
13. Garcia-Losa, M., Unda, M., Badia, X., Rodriguez-Alcantara, F., Carballido, J., Dal-Re, R., & Herdman, M. (2001). ESECI-98 Study Group. Effect of mode of administration on I-PSS scores in a large BPH patient population. *European Urology, 40*, 451–457.
14. Schunemann, H. J., Goldstein, R., Mador, M. J., McKim, D., Stahl, E., Puhan, M., et al. (2005). A randomised trial to evaluate the self-administered standardised chronic respiratory questionnaire. *European Respiratory Journal, 25*, 31–40.
15. Hahn, E. A., Rao, D., Cella, D., & Choi, S. W. (2008). Comparability of interview- and self-administration of the Functional Assessment of Cancer Therapy-General (FACT-G) in English- and Spanish-speaking ambulatory cancer patients. *Medical Care, 46*, 423–431.
16. Wu, A. W., Jacobson, D. L., Berzon, R. A., Revicki, D. A., van der Horst, C., Fichtenbaum, C. J., et al. (1997). The effect of mode of administration on medical outcomes study health ratings and EuroQol scores in AIDS. *Quality of Life Research, 6*, 3–10.
17. Plante, M., Corcos, J., Gregoire, I., Belanger, M. F., Brock, G., & Rossingol, M. (1996). The international prostate symptom score: Physician versus self-administration in the quantification of symptomatology. *Urology, 47*, 326–328.
18. Evans, R. A., Singh, S. J., Williams, J. E., & Morgan, M. D. (2011). The development of a self-reported version of the chronic heart questionnaire. *Journal of Cardiopulmonary Rehabilitation and Prevention, 31*, 365–372.
19. Leung, K. F., Wong, W. W., Tay, M. S., Chu, M. M., & Ng, S. S. (2005). Development and validation of the interview version of the Hong Kong Chinese WHOQOL-BREF. *Quality of Life Research, 14*, 1413–1419.
20. Hays, R. D., Anderson, R., & Revicki, D. (1993). Psychometric considerations in evaluating health-related quality of life measures. *Qual Life Res, 2*, 441–449.
21. Meguro, M., Barley, E., Spencer, S., & Jones, P. (2007). Development and validation of an improved, Respiratory Questionnaire COPD-specific version of the St. George Respiratory Questionnaire. *Chest, 132*, 456–463.

22. Garrod, R., Bestall, J. C., Paul, E. A., Wedzicha, J. A., & Jones, P. W. (2000). Development and validation of a standardized measure of activity of daily living in patients with severe COPD: the London Chest Activity of Daily Living scale (LCADL). *Respiratory Medicine, 94*, 589–596.
23. Ferrer, M., Alonso, J., Prieto, L., Plaza, V., Monso, E., Marrades, R., et al. (1996). Validity and reliability of the St George's Respiratory Questionnaire after adaptation to a different language and culture: the Spanish example. *European Respiratory Journal, 9*, 1160–1166.
24. Vilaro, J., Gimeno, E., Sanchez Ferez, N., Hernando, C., Diaz, I., Ferrer, M., et al. (2007). Actividades de la vida diaria en pacientes con enfermedad pulmonar obstructiva cronica: validacion de la traduccion espanola y analisis comparativo de 2 cuestionarios. *Medicina Clínica (Barc), 129*, 326–332.
25. Bestall, J. C., Paul, E. A., Garrod, R., Garnham, R., Jones, P. W., & Wedzicha, J. A. (1999). Usefulness of the medical research council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax, 54*, 581–586.
26. American thoracic society official statement. Standardization of spirometry. (1994). Update. *American Journal of Respiratory and Critical Care Medicine, 1995*(152), 1107–1136.
27. Scientific Advisory Committee of the Medical Outcomes Trust. (2002). Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research, 11*, 193–205.
28. Ringbaek, T., Martinez, G., & Lange, P. (2012). A comparison of the assessment of quality of life with CAT, CCQ, and SGRQ in COPD patients participating in pulmonary rehabilitation. *Chronic Obstructive Pulmonary Disease, 9*, 12–15.
29. Hood, K., Robling, M., Ingledew, D., Gillespie, D., Greene, G., Ivins, R., et al. (2012). Mode of data elicitation, acquisition and response to surveys: A systematic review. *Health Technology Assessment, 16*, 1–162.
30. Gundy, C. M., & Aaronson, N. K. (2010). Effects of mode of administration (MOA) on the measurement properties of the EORTC QLQ-C30: A randomized study. *Health Qual Life Outcomes., 30*(8), 35. doi:10.1186/1477-7525-8-35.
31. Miravitlles, M., Soriano, J. B., García-Río, F., Muñoz, L., Duran-Tauleria, E., Sanchez, G., et al. (2009). Prevalence of COPD in Spain: impact of undiagnosed COPD on quality of life and daily life activities. *Thorax, 64*, 863–868.