Validation of Polish language version of CAT questionnaire

Authors: Małgorzata Farnik, Grzegorz Brożek, Małgorzata Czajkowska-Malinowska, Rafał Krenke, Aleksander Kania, Marzena Trzaska-Sobczak, Beata Połtyn, Marta Miszczuk, Natalia Celejewska-Wójcik, Krzysztof Kuziemski, Adam Barczyk

Article type: Original article

Received: June 16, 2019.

Accepted: August 10, 2019.

Published online: August 16, 2019.

ISSN: 1897-9483

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY-NC-SA 4.0), allowing third parties to copy and redistribute the material in any medium or format and to remix, transform, and build upon the material, provided the original work is properly cited, distributed under the same license, and used for noncommercial purposes only. For commercial use, please contact the journal office at pamw@mp.pl.
Validation of Polish language version of CAT questionnaire

Małgorzata Farnik ¹, Grzegorz Brożek ², Małgorzata Czajkowska – Malinowska ³⁴, Rafał Krenke ⁵, Aleksander Kania ⁶, Marzena Trzaska – Sobczak ¹, Beata Połtyn ³, Marta Miszczuk ⁵, Natalia Celejewska – Wójcik ⁶, Krzysztof Kuziemski ⁷, Adam Barczyk ¹

1. Department of Pneumonology, School of Medicine in Katowice, Medical University of Silesia, Poland
2. Department of Epidemiology, School of Medicine in Katowice, Medical University of Silesia, Poland
3. Department of Lung Diseases and Respiratory Failure, Kujawy-Pomorze Regional Centre of Pulmonology in Bydgoszcz, Poland
4. Centre of COPD and Respiratory Failure, Kujawy-Pomorze Regional Centre of Pulmonology in Bydgoszcz, Poland
5. Department of Internal Medicine, Pulmonary Diseases and Allergy Medical University of Warsaw, Poland
6. Department of Pulmonology, II Chair of Internal Medicine, Jagiellonian University Medical College, Kraków, Poland
7. Department of Pulmonology and Allergology, Faculty of Medicine, Medical University of Gdańsk, Poland

Short title: Validation of Polish language version of CAT.

Correspondence: Małgorzata Farnik

Department of Pneumonology, School of Medicine in Katowice, Medical University of Silesia, Poland

Ul. Ziołowa 45-47

40-635 Katowice Poland
Email: mfarnik@interia.pl

Conflict of interest: none declared
Abstract

Introduction

Chronic Obstructive Pulmonary Disease Assessment Test (CAT) is a standardized patient–completed tool dedicated to symptoms severity assessment. While a Polish CAT version has been used for a few years, a validation has not been previously completed.

Study objectives

The aim of the study was to validate the Polish-language version of the CAT questionnaire by assessing its reproducibility and reliability.

Patients and Methods

Validation of Polish language version of CAT questionnaire was an substudy of POPE - international, multicenter, observational cross-sectional survey of COPD patients in Central and Eastern European countries. The study was completed in 395 COPD patients in an outpatient settings, stable at least 4 weeks prior to the survey. The validation procedure incorporated statistical methods: Spearman’s correlation and Cronbach’s alpha coefficients, Cohen’s kappa test and Bland–Altman procedure. Statistical significance criterion was established P<0.05.

Results

The internal consistency assessed by Cronbach’s alpha coefficient was 0.87 for the questionnaire and 0.84-0.86 for its separate items. The repeatability of the questionnaire was good to very good (Cohen kappa: 0.76 to 0.85; P<0.01). The Spearman coefficient for the sum of scores of test-retest responses was 0.95 (P<0.01). The Bland and Altman analysis revealed very good test-retest reliability and inter-rater reliability, with the mean difference between test I and test II results -0.556 and 95% CI between -0.345 and +0.767.
Conclusions

The Polish version of the CAT is a reproducible and reliable instrument for evaluation of COPD patients. The version of the questionnaire should be recommended for clinical practice.

Key words: CAT, COPD, reliability, reproducibility, validation
Introduction

Chronic obstructive pulmonary disease (COPD) represents a significant cause of mortality and morbidity throughout the world [1]. According to the Global Initiative for Chronic Obstructive Pulmonary Disease (GOLD) strategy, the diagnosis of COPD is still based on spirometry but symptom assessment is crucial for proper treatment decisions [2]. From a clinician perspective, there is a need for reliable and validated tool for assessment which can improve a communication between the physician and patient. The COPD Assessment Test (CAT) is recommended by GOLD report as a short, simple questionnaire to measure the health status impairment in COPD [3]. The questionnaire comprises eight items relating to cough severity, phlegm, chest tightness, breathlessness, activity limitation, confidence, sleep and energy. The CAT total score ranges from 0 to 40, with score 0 representing no impairment [3]. The tool allows a comprehensive symptom assessment and its importance is growing according to the current GOLD report. Thus CAT has been widely used for assessing and monitoring COPD and several language versions of the CAT questionnaire have been published and validated [5, 6, 7]. The questionnaire is also available in Polish language version (see Supplementary material) and since 10 years widely used in clinical practice but the validation of this tool in Poland has not been performed before. The aim of the study was to validate the Polish-language version of the CAT questionnaire by assessing its reproducibility and reliability, as only validated questionnaires should be recommended for research and clinical practice.

Validation of the original English language version of the questionnaire has been completed previously and it was found that the tool had good measurement properties with excellent internal consistency [3]. The Cronbach’s alpha coefficient for the questionnaire was very good (alpha=0.88; P<0.05). The validation process of original version of CAT
questionnaire included intra-class correlation, as well as sensitivity to change [3]. The original version was validated for specific items, in addition selected clinical criteria were incorporated.

**Patients and Methods**

**Study design**

The study was a substudy of the POPE project – an international, multicenter, observational cross-sectional survey of COPD patients in Central and Eastern European countries. During the study the validation process of CAT questionnaire has been completed involving different sites representing a range of Polish regions which were participating in the POPE study [7, 8]. The coordinating center was Medical University of Silesia in Katowice. The study was performed in accordance with the ethics principles of the Declaration of Helsinki. The Ethics Committee of the Polish coordinating center, after reviewing the study protocol, had decided that the ethical approval was not needed due to a non-interventional study design.

The methodology of the POPE study was described in detail elsewhere [7] Patients with clinically confirmed COPD diagnosis at least 12 months before the visit day were recruited to the study. The diagnosis was based on clinical data and irreversible airway obstruction in spirometry. Inclusion criteria were: age 40-80 years and stable course of disease for at least 4 weeks prior to the survey. Smoking history was not obligatory, other exposure risk factors were also allowed. Comorbid diseases were allowed, but exacerbation of any medical condition was construed as an exclusion criterion.

Validation of newly developed measure requires more complex statistical analysis. If original questionnaire is translated into other language version the literal or linguistic equivalence of both language versions do not ensure that they have the same psychometric properties. The
consistency of translated version of the questionnaire can be evaluated using its internal consistency, test re-test validity [9].

The validation procedure of translated version Polish language version of CAT questionnaire included complex methods applied in the reliability analysis [10]

- the analysis of statistical properties of test items (the assessment of internal consistency) and the relation of items to general test result
- test-retest reliability: the comparison of double tests with the same method (the estimation of the internal stability of the test).

The Polish version of CAT questionnaire was self-completed by a patient during the visit in the outpatient pulmonary department. To assess test-retest reliability the questionnaire was completed twice by each patient with one hour between completions. There is no specific limitation for the period of time required between each completion of the questionnaire in the test-retest reliability procedure. Symptoms or complains are expected not to change significantly during one hour in the case of stable COPD patient. Period of time is also long enough in the context of reminding all answers chosen during previous completion.

The language version questionnaire which was validated during the study was translated into Polish by GOLD Committee affiliates. The translation process was based on the standard procedures of questionnaires translation [10].

**Internal consistency**

Testing for homogeneity of the measurement is an important procedure assessing the reliability of a the instrument. Internal consistency is defined as the correlations between the items in the scale or within each scale domain or correlations between the items and the total score. Internal consistency is measured by applying Cronbach’s alpha coefficient, the
calculation is based on average correlation among the items and the number of items in the instrument, thus the coefficient is reaching the values between 0 and 1 [11, 12].

**External validity**

In order to assess the external validity, the analysis also included an assessment of the association between CAT scores and the mMRC dyspnea scale, and spirometry parameters: pre and postbronchodilator spirometry (post-BD FEV1 [l] and [% of predicted], post-BD FVC [l] and [% of predicted], post-BD FEV1 %FVC) as well as the 6 Minutes Walking Test (6MWT).

**Test-retest reliability**

Reliable measure means that it is stable or consistent and produces similar results when administrated repeatedly, when it is no evidence of change. To assess test – retest reliability the instrument is administered to the same population on two occasions (stable over the interval between assessments) and the two scores are assessed for consistency. The results could be influenced by the possibility of practice effects, which can artificially inflate the estimate of reliability [13, 14].

The repeatability of the questionnaire was assessed using the Bland – Altman procedure and the Cohen’s kappa statistical test.

**Statistical analyses**

The statistical analysis was performed using standard procedures available in the Statistica 12.0 software package (Stat-Soft Inc, USA) and SAS, version 9.2 (SAS Institute Inc.,Gary, NC). Normality of distributions of continuous variables was assessed by the Shapiro-Wilk test. Statistical significance of differences between continuous variables were analyzed by the Student’s t-test and if non-normal distribution was found - U Mann-Whitney
test and Wilcoxon signed-rank test for paired variables. Differences between categorical variables were examined by the Chi-square test.

A level of internal consistency of the test was assessed by analyzing the correlation of answers to questions with a total score of the questionnaire on the basis of Cronbach’s statistics. The raw and standardized Cronbach’s α coefficients were calculated (value scaling of variables – answers to questions with the assumption that the standard deviation equals 1). A satisfactory level of consistency was defined by the standardized value of α statistics > 0.70. Apart from calculating general α statistics, the impact of separate questions on the consistency level of the questionnaire was defined by analyzing a potential improvement of the α value after a possible removal of subsequent questions from the questionnaire.

In order to assess the external accuracy of the tool, the analysis also included an assessment of the association between CAT scores and the mMRC dyspnea scale, 6MWT, and spirometry parameters.

Estimation of the repeatability of the responses was conducted using the Bland – Altman procedure as well as the Cohen’s kappa statistical test. In the study, the statistical agreement was determined with the conventional scale where the kappa values have the following meaning: 0.81–1.00 almost perfect agreement, 0.61–0.80 substantial agreement, 0.41–0.60 moderate agreement, 0.20–0.40 fair agreement, < 0.20 slight agreement [15].

The statistical inferences were based on the level of significance of P < 0.05.

Results

Study group characteristics

The study group comprised 395 ambulatory COPD patients (258 males and 37 females). The mean (SD) age of patients was 67.9 (9.7) years with no statistically significant differences between gender groups (P ≥ 0.1). Most of the patients were ex-smokers 73.67%,
current smokers 20.76%, passive smokers 2.28% and only 3.29% having non-smoking history. The only difference between males and females was smoking history with males representing significantly higher exposure measured as pack years (Table 1).

**Internal consistency reliability**

The reliability analysis based on the baseline completion of the questionnaire, showed the Cronbach’s α raw coefficient 0.87, and the standardized coefficient 0.86. Table 2 presents the impact of particular questions on the general level of consistency of the test defined by the values of alpha coefficients were ranging from 0.84 to 0.86. The results of partial correlations analysis showed that all correlations reached a value of 0.5 or higher and ranged from 0.5 to 0.74. Table 2 presents the correlation coefficients between the responses to particular questions and a general result of the test and the values of the Cronbach’s α raw and standardised coefficients after a possible removal of a given question from the questionnaire.

Correlations between answers for an individual questions were statistically significant (P<0.05) and reached expected values – ranging from 0.30 to 0.74 (Table 3).

**Test-retest reliability and inter-rater reliability**

To assess test-retest reliability, the instrument was administered to the same study group again after one hour. There were no differences (P>0.05) between the results for single questions or for the total scores obtained during the two measurements. The correlations between the total scores from test-retest measurements were very good (Spearman rank correlation R=0.95; P<0.001). The Polish language CAT questionnaire was characterized by a very good repeatability with Kappa coefficient ranging from 0.76 to 0.85 (P<0.01). The details are described in Table 4.
According to Wilcoxon signed-rank test for paired variables there were no difference (P>0.05) between results for single questions as well as for the total scores obtained during the two measurements.

The Bland and Altman analysis also revealed very good test-retest reliability and inter-rater reliability, with the mean difference between test I and test II results -0.556 and 95% CI between -0.345 and +0.767 (Figure 1.)

The relation between the CAT and other clinical measures

The analysis of external validity of the CAT scores had shown significant (P<0.05) correlations between the Polish language version of the CAT total score and mMRC (r=-0.57), 6MWT (r=-0.32) and some pulmonary function measurement, like FEV1 in liters (r=-0.37), FEV1 % of predicted value (r=-0.38), FVC in liters (r=-0.12), FVC % of predicted value (r=-0.31) and FEV1/FVC (r=-0.22) (Table 5).

The CAT scores were not conditioned by participants’ age or gender as well as education level.

Discussion

The results of our study showed that the psychometric properties of the Polish version of the questionnaire were satisfactory. The validation process included the assessment of internal consistency, the relation of test items with the general test result and test-retest reliability as the estimation of the internal stability of the test.
CAT questionnaire has been translated into several languages and some of them has been validated. Study results are consistent with those previously reported in other populations, from both Europe [16] and outside of Europe [4, 17]. Our study showed the expected reliability for each item with the Cronbach’s alpha coefficient from 0.83 to 0.86 for individual items. The Cronbach’s alpha coefficient for the questionnaire was 0.87, thus could be interpreted as very good and was similar as for the original version (alpha=0.88) [3]. For example the Crombach’s alpha coefficient for Hindi was 0.83 [4] and 0.85 for Korean [5].

Test – retest of the Polish language version was very good with Spearman rank coefficient R=0.95 (P<0.001) with good repeatability also for each question. The original version of questionnaire test-retest was also very good with an internal consistency correlation coefficient 0.8 [3].

The validation of original version didn’t include the relation between questionnaire results and other clinical parameters such as dyspnea scale or spirometry. In our study, the relation of test items with the other clinical data has been analyzed. The instruments we used in our study were the mMRC dyspnea scale, 6MWT, and some spirometry parameters. We have observed significant correlations between the Polish language CAT total score and mMRC, 6MWT, FEV₁, FVC and FEV₁/FVC. In our study, similarly to other observations, total scores of the questionnaire correlated significantly with mMRC dyspnea scale, which confirms that there was a relation activity limitation due to dyspnea and CAT results [6, 18, 19]. The correlation analysis showed weak correlation between dyspnea and CAT total scores. The relatively low correlation between CAT scores and mMRC scale suggests using both tools for more complex clinical assessment. CAT questionnaire concerns more wide spectrum of patient’ everyday functioning, and questions which are focused on dyspnea is only a part of assessment. Despite CAT and mMRC are proposed as equivalent measures for patients classification and
treatment stratification [2] the range of correlation between both measures reported in other studies varies from 0.29 to 0.62 [20, 21].

In case of 6MWT and spirometry parameters we observed weak but significant inverse correlations with CAT score. It confirms that in our study airflow limitation and limited physical activity were accompanied by increasing CAT scores. Similar negative correlation between 6MWT results and CAT scores were found in Portuguese study [6]. The correlation between CAT scores and spirometry were analyzed only in some of the language validations as additional observations. The validation of original version didn’t discuss this issue. The relation between CAT results and FEV1 and FVC was raised in the validation into Arabic and was found as not significant [22], but validation of Hindi [5] and Portuguese language versions had shown significant correlations [6]. Thus most of our results were in line with the observations from others studies [5, 6, 17, 18].

Study limitations

The hypothetic limitation in our study is that the sensitivity to change was not assessed during the study. Sensitivity to change is defined by instrument’s responsiveness to detect the change. It requires correlating its scores with other measures which reflect any anticipated changes. The responsiveness is not required if the validation process concerns translated version of originally validated questionnaire; thus this issue was not raised during our validation study as well as in other language validation studies, but seems to be crucial for the assessment of questionnaire utility. Additionally, the responsiveness was not done in validation process of original CAT validation study.

Strengths of the study

The notable strength of the study was relatively large, community-based, urban- rural study population, as Polish CAT validation study was substudy of multicenter study – POPE.
Patients had represented different age and educational level, all participants were recruited based on real-life criteria of study, thus data could be translated into everyday practice.

Apart from good data quality, several independent statistical methods suitable for test validation were used and relation between CAT results and clinical data had been analyzed.

Significant heterogeneity of COPD patients due to various clinical presentation, response to therapy, as well as concomitant diseases has an impact of patients’ functioning. Our study showed, similar to other studies, that age [23] or concomitant diseases [24] are related with disease severity. Thus symptom assessment is crucial in the context of treatment and should be based on simple and reliable methods. Patient related outcome measures should correspond to the specific clinical situation and bring an opportunity to improve the quality of care. The Polish version of the questionnaire CAT meets all necessary criteria expected during a validation process and could be recommended for clinical practice. Despite this, there are still some issues on which we need to focus our attention in further research. So far CAT is one of the available tools to establish a threshold at which patients become sufficiently symptomatic to justify regular treatment. Further studies are required in the context of this tool utility for disease. Consistent use of the same validated methods of such assessments would allow us to create a more patient - orientated approach in COPD treatment.

Conclusions

The Polish language version of the CAT questionnaire is a valid, reproducible, and reliable instrument for evaluating patients with COPD. We conclude, that this version of the questionnaire should be recommended for clinical assessment practice in Poland.

Contribution

Contribution
MF Concept of the study; Writing – original draft Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing; GB Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing; MCM Formal analysis, Investigation; RK Methodology, Formal analysis, Investigation; AK Methodology, Formal analysis, Investigation; MTS Formal analysis, Investigation, writing- review; BP Investigation; MM Investigation, NCW Methodology, Investigation; KK Methodology, Investigation; AB, Writing – original draft, Supervision.

Acknowledgments

The study team would like to thank GlaxoSmithKline affiliates for kind validation process permission. CAT is a trademark GlaxoSmithKline group. Authors would like to thank also the Ludwig Boltzman Institute for COPD and Respiratory Epidemiology in (Vienna, Austria) – the sponsor of the POPE study (ClinicalTrials.gov. registry identifier: NCT02119494). This research institute received unrestricted grant from Boehringer Ingelheim.

Special thanks for Joshua Lawson for review of the manuscript and linguistic remarks.
References


# Table 1. Study group characteristics

<table>
<thead>
<tr>
<th></th>
<th>Total Median (Q1; Q3)</th>
<th>Males Median (Q1; Q3)</th>
<th>Females Median (Q1; Q3)</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of patients, y-rs</td>
<td>67.00 (61.00; 75.00)</td>
<td>69.0 (60.0; 77.0)</td>
<td>66.0 (61.0; 73.0)</td>
<td>0.1</td>
</tr>
<tr>
<td>Illness duration, y-rs</td>
<td>8.00 (3.00; 12.00)</td>
<td>9.00 (4.00; 13.00)</td>
<td>7.00 (2.00-12.00)</td>
<td>0.09</td>
</tr>
<tr>
<td>Age at diagnosis, y-rs</td>
<td>59.00 (52.00; 67.00)</td>
<td>60.00 (52.00; 67.00)</td>
<td>59.00 (53.00; 65.00)</td>
<td>0.6</td>
</tr>
<tr>
<td>Smoking, pack years</td>
<td>36.00 (23.00; 50.00)</td>
<td>39.50 (25.00; 54.00)</td>
<td>30.00 (20.00; 44.00)</td>
<td>0.0004</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>27.59 (24.00; 30.90)</td>
<td>27.57 (24.31; 31.24)</td>
<td>27.59 (23.07; 30.43)</td>
<td>0.2</td>
</tr>
<tr>
<td>6MWD, m</td>
<td>400.00 (327.00; 460.00)</td>
<td>400.00 (330.00; 460.00)</td>
<td>399.00 (322.00; 450.00)</td>
<td>0.9</td>
</tr>
<tr>
<td>mMRC</td>
<td>2.00 (1.00; 3.00)</td>
<td>2.00 (1.00; 3.00)</td>
<td>2.00 (1.00; 3.00)</td>
<td>0.9</td>
</tr>
<tr>
<td>Post-BD -FEV₁, % pred.</td>
<td>52.40 (39.48; 68.37)</td>
<td>51.56 (38.24; 67.90)</td>
<td>55.09 (44.39; 68.89)</td>
<td>0.07</td>
</tr>
<tr>
<td>FEV₁/FVC</td>
<td>0.50 (0.38; 0.60)</td>
<td>0.47 (0.38; 0.59)</td>
<td>0.53 (0.40; 0.62)</td>
<td>0.02</td>
</tr>
</tbody>
</table>

* result of U Mann-Whitney test; Abbreviations: Q1 - first quartile, Q3- third quartile, BMI, Body Mass Index; FEV₁, Forced Expiratory Volume in 1 second; FVC; Forced Vital Capacity; m, meters; mMRC – modified Medical Research Council scale; 6MWT, 6 Minute Walking Test; Post-BD, post bronchodilator; pred, predicted; SD, standard deviation; y-rs, years; m, meters
Table 2. Impact of the responses to particular questions of the Polish language version of CAT questionnaire on questionnaire’s reliability.

<table>
<thead>
<tr>
<th>Question</th>
<th>Correlation coefficient</th>
<th>Cronbach’s alpha coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Raw</td>
</tr>
<tr>
<td>1</td>
<td>0.50</td>
<td>0.86</td>
</tr>
<tr>
<td>2</td>
<td>0.54</td>
<td>0.86</td>
</tr>
<tr>
<td>3</td>
<td>0.52</td>
<td>0.86</td>
</tr>
<tr>
<td>4</td>
<td>0.67</td>
<td>0.84</td>
</tr>
<tr>
<td>5</td>
<td>0.72</td>
<td>0.84</td>
</tr>
<tr>
<td>6</td>
<td>0.74</td>
<td>0.84</td>
</tr>
<tr>
<td>7</td>
<td>0.61</td>
<td>0.85</td>
</tr>
<tr>
<td>8</td>
<td>0.68</td>
<td>0.85</td>
</tr>
</tbody>
</table>
Table 3. Spearman correlation coefficients between the answers for an individual CAT questions.

<table>
<thead>
<tr>
<th>No of question</th>
<th>Question 1</th>
<th>Question 2</th>
<th>Question 3</th>
<th>Question 4</th>
<th>Question 5</th>
<th>Question 6</th>
<th>Question 7</th>
<th>Question 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 1</td>
<td>1.00</td>
<td>0.62</td>
<td>0.30</td>
<td>0.31</td>
<td>0.32</td>
<td>0.34</td>
<td>0.33</td>
<td>0.37</td>
</tr>
<tr>
<td>Question 2</td>
<td>0.62</td>
<td>1.00</td>
<td>0.36</td>
<td>0.37</td>
<td>0.34</td>
<td>0.40</td>
<td>0.39</td>
<td>0.34</td>
</tr>
<tr>
<td>Question 3</td>
<td>0.30</td>
<td>0.36</td>
<td>1.00</td>
<td>0.38</td>
<td>0.37</td>
<td>0.45</td>
<td>0.46</td>
<td>0.33</td>
</tr>
<tr>
<td>Question 4</td>
<td>0.31</td>
<td>0.37</td>
<td>0.38</td>
<td>1.00</td>
<td>0.74</td>
<td>0.62</td>
<td>0.44</td>
<td>0.56</td>
</tr>
<tr>
<td>Question 5</td>
<td>0.32</td>
<td>0.34</td>
<td>0.37</td>
<td>0.74</td>
<td>1.00</td>
<td>0.70</td>
<td>0.45</td>
<td>0.65</td>
</tr>
<tr>
<td>Question 6</td>
<td>0.34</td>
<td>0.40</td>
<td>0.45</td>
<td>0.62</td>
<td>0.70</td>
<td>1.00</td>
<td>0.54</td>
<td>0.65</td>
</tr>
<tr>
<td>Question 7</td>
<td>0.33</td>
<td>0.39</td>
<td>0.46</td>
<td>0.44</td>
<td>0.45</td>
<td>0.54</td>
<td>1.00</td>
<td>0.50</td>
</tr>
<tr>
<td>Question 8</td>
<td>0.37</td>
<td>0.34</td>
<td>0.33</td>
<td>0.56</td>
<td>0.65</td>
<td>0.65</td>
<td>0.50</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Table 4. Kappa coefficient for each question CAT I and CAT II (n=322).

<table>
<thead>
<tr>
<th>Question</th>
<th>Measurement I Mean [SD]</th>
<th>Measurement II Mean [SD]</th>
<th>P*</th>
<th>Kappa coefficient (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 1</td>
<td>2.34 [1.22]</td>
<td>2.33 [1.69]</td>
<td>0.8</td>
<td>0.84 (0.79-0.88)</td>
</tr>
<tr>
<td>Question 2</td>
<td>2.29 [1.37]</td>
<td>2.26 [1.32]</td>
<td>0.7</td>
<td>0.77 (0.72-0.81)</td>
</tr>
<tr>
<td>Question 3</td>
<td>1.73 [1.42]</td>
<td>1.79 [1.43]</td>
<td>0.6</td>
<td>0.77 (0.72-0.82)</td>
</tr>
<tr>
<td>Question 4</td>
<td>3.57 [1.47]</td>
<td>3.62 [1.43]</td>
<td>0.3</td>
<td>0.85 (0.81-0.88)</td>
</tr>
<tr>
<td>Question 5</td>
<td>2.83 [1.60]</td>
<td>2.94 [1.58]</td>
<td>0.6</td>
<td>0.82 (0.78-0.86)</td>
</tr>
<tr>
<td>Question 6</td>
<td>2.08 [1.57]</td>
<td>2.16 [1.59]</td>
<td>0.5</td>
<td>0.82 (0.78-0.86)</td>
</tr>
<tr>
<td>Question 7</td>
<td>2.01 [1.59]</td>
<td>2.04 [1.59]</td>
<td>0.7</td>
<td>0.81 (0.76-0.85)</td>
</tr>
<tr>
<td>Question 8</td>
<td>2.58 [1.49]</td>
<td>2.63 [1.45]</td>
<td>0.7</td>
<td>0.76 (0.71-0.81)</td>
</tr>
<tr>
<td>Total</td>
<td>19.44 [8.54]</td>
<td>19.78 [1.02]</td>
<td>0.6</td>
<td>0.81 (0.77-0.85)</td>
</tr>
</tbody>
</table>

* difference between two CAT; measurements by Wilcoxon signed-rank test;

Abbreviations: 95%CI, 95% Confidence Interval; SD, standard deviation
Table 5. Spearman’s rank correlations between Polish language version of CAT and modified Medical Research Council dyspnea scale, six-minute walk test as well as selected spirometry values.

<table>
<thead>
<tr>
<th>Variable</th>
<th>R</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>mMRC</td>
<td>0.57</td>
<td>0.03</td>
</tr>
<tr>
<td>6MWT</td>
<td>-0.32</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FEV₁ [l]</td>
<td>-0.37</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FEV₁ [% of predicted]</td>
<td>-0.38</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FVC [l]</td>
<td>-0.12</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FVC [% of predicted]</td>
<td>-0.31</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FEV₁/FVC</td>
<td>-0.22</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Abbreviations: FEV₁, Forced Expiratory Volume in 1 second; FVC; Forced Vital Capacity; l, liters; mMRC – modified Medical Research Council scale; 6MWT, 6 Minute Walking Test;
Figure 1. Repeatability of the measurement of the Polish version of CAT – the result of the Bland-Altman procedure.

Supplementary material

Polish-language version of the COPD Assessment Test