

ORIGINAL ARTICLE

Baseline health status and setting impacted minimal clinically important differences in COPD: an exploratory study

Harma Alma^{a,b,*}, Corina de Jong^{a,b}, Danijel Jelusic^c, Michael Wittmann^c, Michael Schuler^d,
Boudewijn Kollen^a, Robbert Sanderma^{e,f}, Janwillem Kocks^{a,b}, Konrad Schultz^c,
Thys van der Molen^{a,b}

^aDepartment of General Practice and Elderly Care Medicine, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

^bGroningen Research Institute for Asthma and COPD (GRIAC), University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

^cKlinik Bad Reichenhall, Center for Rehabilitation, Pulmonology and Orthopedics, Bad Reichenhall, Germany

^dInstitute for Clinical Epidemiology and Biometry (ICE-B), Julius-Maximilians-Universität Würzburg, Würzburg, Bayern, Germany

^eDepartment of Health Psychology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

^fDepartment of Psychology, Health and Technology, University of Twente, Enschede, The Netherlands

Accepted 23 July 2019; Published online 27 July 2019

Abstract

Objectives: Minimal clinically important differences (MCIDs) are used as fixed numbers in the interpretation of clinical trials. Little is known about its dynamics. This study aims to explore the impact of baseline score, study setting, and patient characteristics on health status MCIDs in chronic obstructive pulmonary disease (COPD).

Study Design and Setting: Baseline and follow-up data on the COPD Assessment Test (CAT), Clinical COPD Questionnaire (CCQ), and St. George's Respiratory Questionnaire (SGRQ) were retrospectively analyzed from pulmonary rehabilitation (PR) and routine clinical practice (RCP). Anchor- and distribution-based MCID estimates were calculated and tested between settings, gender, age, Global initiative

Funding: The main RIMTCORE trial (#DRKS00004609) was funded by the Deutsche Rentenversicherung, Germany. The Dutch observational study on COPD health status in routine clinical practice (MCID study) as well as the current combined retrospective analysis of both studies received financial support from the Junior Scientific Masterclass (JSM) as part of the University of Groningen, the Netherlands.

Conflict of interest: H.J. Alma, C. de Jong, D. Jelusic, M. Wittmann, M. Schuler, B.J. Kollen, and R. Sanderma have nothing to disclose. J.W.H. Kocks reports personal fees from Novartis; research grants and personal fees from Boehringer Ingelheim; research grants and personal fees from GlaxoSmithKline (GSK); research grants from Stichting Zorgdraad; personal fees from the International Primary Care Respiratory Group (IPCRG); personal fees from Springer Media; and travel arrangements from Chiesi Pharmaceuticals BV, GSK BV, and IPCRG, all outside the submitted work. K. Schultz received lecture fees from Boehringer, Astra-Zeneca, Berlin Chemie, Novartis, Chiesi, Mundipharma, Takeda, GSK, and MSD, all outside the submitted work. T. van der Molen reports personal reimbursements from GSK, TEVA, Astra Zeneca, and Boehringer Ingelheim and study grants from Astra Zeneca and GSK. After this study was terminated, he became an employee of GSK. None of these prior stated conflicts of interest are linked to the current manuscript. T. van der Molen developed the CCQ and holds the copyright.

Ethics approval and consent to participate: This retrospective study is a secondary analysis of a subsample from the Routine Inspiratory Muscle Training within COPD Rehabilitation (RIMTCORE) real-life randomized controlled trial (#DRKS00004609) in the Klinik Bad Reichenhall, Center for Rehabilitation, Pulmonology, and Orthopaedics in Germany and a primary analysis of all patients participating in the Dutch observational study on COPD health status in routine clinical practice (MCID Study; UMCG trial #201500447). All patients in both studies signed informed consent on participation. The

RIMTCORE trial has been approved by the Ethik-Kommission der Bayerischen Landesärztekammer and registered in the German Clinical Trial Register. The MCID study has been registered at the University Medical Center Groningen (UMCG) Research Register and evaluated by its Medical Ethical Committee.

Authors' contributions: K.S., M.W., D.J. and M.S. planned the RIMTCORE study design and were responsible for data collection. H.A., C.d.J., R.S., and T.v.d.M. designed the Dutch observational study on COPD health status in routine clinical practice (MCID study) as well as the current retrospective analysis of both studies. H.A., C.d.J., and B.K. performed the statistical analysis. H.A. wrote the first draft, whereas C.d.J., B.K., J.K., R.S., and T.v.d.M. actively participated in the review process. R.S. and T.v.d.M. supervised and participated in different steps of the study, as well as in writing. All authors participated in and approved the final version of the manuscript.

Consent of publication: All authors participated in various steps in the study, edited the manuscript, and gave their approval for submission.

Data sharing statement: The data that support the findings of this study are not publicly available. Participating patients in the RIMTCORE trial have only agreed on availability of their data to the Klinik Bad Reichenhall, their scientific partners in the data analysis, and the Committee of the Bavarian State Chamber of Labor in Munich. Participating patients in the MCID study only agreed on availability of their data to the University Medical Center Groningen (UMCG) and their scientific partners in the data analysis.

* Corresponding author. Department of General Practice and Elderly Care Medicine, University of Groningen, University Medical Center Groningen, HPC FA21, P.O. Box 196, NL-9700 AD, Groningen, The Netherlands. Tel.: +31 6 27137024.

E-mail address: h.j.alma@umcg.nl (H. Alma).

for Obstructive Lung Disease (GOLD) classification, comorbidities, and baseline health status.

Results: In total, 658 patients were included with 2,299 change score measurements. MCID estimates for improvement and deterioration ranged for all subgroups 0.50–6.30 (CAT), 0.10–0.84 (CCQ), and 0.33–12.86 (SGRQ). Larger MCID estimates for improvement and smaller ones for deterioration were noted in patients with worse baseline health status, females, elderly, GOLD I/II patients, and patients with less comorbidities. Estimates from PR were larger.

Conclusion: Baseline health status and setting affected MCID estimates of COPD health status questionnaires. Patterns were observed for gender, age, spirometry classification, and comorbidity levels. These outcomes would advocate the need for tailored MCIDs. © 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Chronic obstructive pulmonary disease (COPD); Health status; Minimal clinically important difference (MCID); COPD Assessment Test (CAT); Clinical COPD Questionnaire (CCQ); St. George's Respiratory Questionnaire (SGRQ)

1. Background

Health status measurement has been defined as “a standardized way of quantifying the impact of disease on a patient's life, health, and well-being” [1]. It provides important information about the burden of disease—especially in patients with chronic obstructive pulmonary disease (COPD)—because physiologic measures like spirometry alone do not measure the full aspect of this chronic disorder [1–7]. Many factors may impact health status in COPD, including the number and severity of exacerbations [8–10]; the disease severity defined according to the Global initiative for Obstructive Lung Disease (GOLD) spirometry classification [5,9–12]; the patient's gender and age [5,8–10,13–15]; and the amount of comorbidities [12]. Moreover, the baseline score may also be predictive of (worsening) health status [11,14].

Health status is, in addition to spirometry, an obligatory endpoint in evaluating treatment outcomes of clinical trials [16]. An intervention or therapy should result in meaningful clinical changes using outcomes that are relevant for patients themselves [16–18]. A parameter to interpret and define important change in health status is the minimal clinically important difference (MCID). It has been defined as “the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in absence of troublesome side effects and excessive costs, a change in patient management” [19]. Like health status, MCIDs may also be influenced by multiple patient- and disease-related factors. Previous publications speculated that MCIDs for quality of life tools may be affected by patient characteristics such as age, gender, and education; the number of comorbidities; initial baseline condition of the patient; and pathologic characteristics of the disorder including disease severity [16,17,20–39]. It remains, however, unclear whether different MCIDs should be used in practice depending on disease severity, patient characteristics, and baseline health state.

Various studies have discussed the MCID of COPD health status tools [40], including the recommended questionnaires COPD Assessment Test (CAT), Clinical COPD

Questionnaire (CCQ), and the St. George's Respiratory Questionnaire (SGRQ) [2,41,42]. The MCIDs were determined as a rather fixed range of two to three points on the CAT [40,43–49], 0.40–0.50 on the CCQ [40,43,44,48–52], and four up to seven points on the SGRQ [40,43,44,53–57]. However, little is known about the dynamics of these MCIDs and the impact of different (disease) characteristics of patients on it. The present study aims to explore patient- and disease-related influences upon the MCID estimates for these recommended instruments in patients with COPD from different settings, in this case pulmonary rehabilitation (PR) and routine clinical practice (RCP).

2. Methods

2.1. Study subjects and design

Data from patients with COPD in two studies were retrospectively analyzed: (1) the 3-week full-day inpatient Routine Inspiratory Muscle Training within COPD Rehabilitation (RIMTCORE) real-life randomized controlled trial during PR in Klinik Bad Reichenhall (Center for Rehabilitation, Pulmonology, and Orthopedics) in Germany [43,58] and (2) an observational trial (MCID study) aimed primarily at measuring health status changes in Dutch primary and secondary RCP [59].

2.2. Data collection

Patient descriptive characteristics and postbronchodilator spirometry were collected at baseline. CAT (no recall), CCQ (1 week recall), and SGRQ (1 month recall) were scored at baseline and during follow-up as study outcomes. In the RIMTCORE trial, follow-up questionnaires were scored at discharge (3 weeks) and at 3/6/9/12 months (Fig. 1). In RCP, respective health status questionnaires were scored at baseline and after 3/6/12 months. The CAT contains eight items and a total maximum score of 40 (maximum impairment) [60]. The CCQ consists of ten items and a total maximum score of six (maximum

What is new?**Key findings**

- Larger minimal clinically important difference (MCID) estimates for improvement and smaller ones for deterioration on chronic obstructive pulmonary disease (COPD) health status tools were observed in patients with worse baseline health status, females, elderly, Global initiative for Obstructive Lung Disease I/II patients, and in patients with less comorbidities.
- MCID estimates for COPD health status tools were generally larger for both improvement and deterioration during a pulmonary rehabilitation intervention compared with routine clinical practice without intervention.

What this adds to what was known?

- This is the first study to explore the dynamics of the MCID of COPD health status tools by integrating the influence of baseline health status, patient characteristics, disease severity, and study setting.

What is the implication and what should change now?

- Currently, fixed MCIDs of COPD health status tools are used for the interpretation of clinical trial results; however, the present study would advocate the need for clustered or even tailored MCIDs to enable a more individual evaluation of change scores.

impairment) [61]. The SGRQ contains 50 items and a total score of 100 (maximum impairment) [62]. A 15-point Global Rating of Change (GRC) scale was scored in both studies for each follow-up moment. It required patients to assess their experienced change in COPD health status compared with baseline. Answers were scored on a scale from minus seven (very much worse) to plus seven (very much better) with zero equaling no change [63].

2.3. Study methods to determine MCIDs

Health status change scores were calculated as the difference between follow-up and baseline scores. Negative change represented improvement, and positive change represented deterioration. MCID estimates were calculated using anchor- and distribution-based methods. The anchor-based approach required patients to be categorized according to their GRC score. Scores of 0 and ± 1 on the GRC scale indicated no change; scores of ± 2 and ± 3 represented minimal improvement/deterioration; scores of ± 4

and ± 5 were considered moderate improvement/deterioration; and scores of ± 6 and ± 7 indicated large improvement/deterioration [63]. The mean health status change score of the minimal change group according to the GRC represented the MCID estimate, assuming normality of distribution. In addition, the GRC score for minimally improved and deteriorated was used as a cutpoint for dichotomization in the receiver operating characteristics (ROC) curve analysis [64]. Finally, the distribution-based approach included the half standard deviation (0.5SD) of the health status change score [65]. Study methods were performed separately for improvement and deterioration. MCID estimates were evaluated and statistically tested between subgroups for study settings (PR and RCP); males and females; low and high age (median as cutpoint); low (GOLD I–II) vs. high (GOLD III–IV) COPD disease severity; low and high levels of comorbidities on the Charlson Index (median as cutpoint); and better and worse baseline health status (median as cutpoint).

2.4. Data analysis and statistics

SPSS, version 25.0, (IBM, Chicago) was used for data analysis. Baseline, follow-up, and change scores were assessed. Frequencies with percentages, mean with SD, or median with interquartile range (IQR) were determined depending on the variable characteristics. Normality of distribution was assessed for continuous variables using histograms, and skewness and kurtosis with values between -1 and $+1$ indicative for normality. Baseline scores between both data sets were tested for a difference using independent *t*-tests, Mann Whitney *U*-tests, or chi-square tests depending on type of data and whether assumptions were met. Health status and GRC scores were assessed for floor and ceiling effects, defined as over 15% of patients scoring in the lowest and highest 10% of the maximum scale range [66]. Baseline and follow-up health status scores were tested for significance of change using paired *t*-tests or Wilcoxon-signed rank tests. Pair-wise deletion was applied. The false discovery rate (FDR) due to multiple testing was controlled for by Benjamini–Hochberg’s step-up procedure to maintain an overall two-sided type I error rate of 5% [67]. Adjusted *P*-values are reported. Intraclass correlation coefficients (ICCs) of the health status change scores were determined to assess dependency of data. The impact of regression to the mean was determined as $100(1-r^{1/2})$ with *r* reflecting the ICC.

Correlations were assessed between the GRC and health status change scores using Pearson coefficients assuming normality of distribution and the GRC to be a continuous variable. Correlations were required ≥ 0.30 [68]. Patients were categorized per GRC class. Significance of change within each GRC category was tested using paired *t*-tests or Wilcoxon-signed rank tests. ANOVA was performed to statistically test for significant differences between GRC categories and to determine the mean change scores

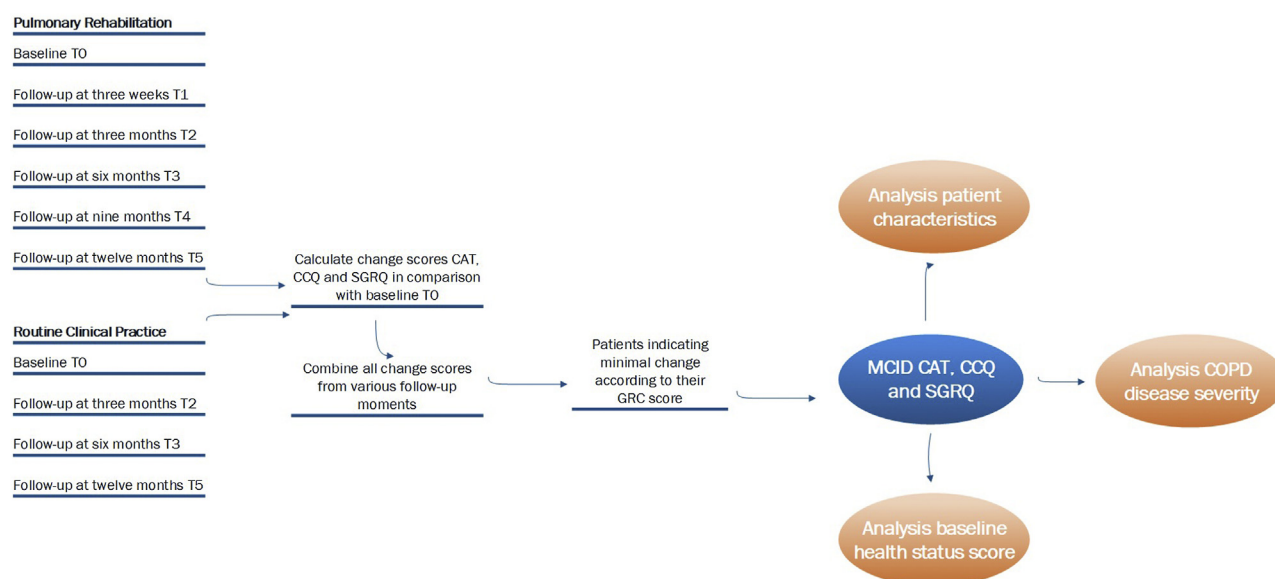


Fig. 1. Study methods of the MCID (subgroup) analysis. Abbreviations: CAT, COPD Assessment Test; CCQ, Clinical COPD Questionnaire; COPD, chronic obstructive pulmonary disease; GRC, Global Rating of Change; SGRQ, St. George's Respiratory Questionnaire; T0, baseline; T1, follow-up at 3 weeks; T2, follow-up at 3 months; T3, follow-up at 6 months; T4, follow-up at 9 months; T5, follow-up at 12 months.

including confidence interval per class. The mean change scores of the minimally improved and/or deteriorated group of patients represented the MCID estimates. Furthermore, ROC curves used a GRC score of ± 2 as cutoff values to dichotomize the data set to differentiate between important and unimportant change [24,34,38]. The area under the curve (AUC) was calculated, and the optimal combination

of sensitivity and specificity was selected as MCID estimate (lowest sum of $[1 - \text{sensitivity}]$ and $[1 - \text{specificity}]$) [64].

Independent *t*-tests were used to determine significance of the difference between MCID estimates for the various subgroups. These subgroups were determined for nominal values as the respective categories. For the continuous

Table 1. Patient characteristics and health status scores

Variable	Pulmonary rehabilitation	Routine clinical practice	Significance
Number of patients (<i>N</i>)	451	207	
Age (years) ^a	57.87 ± 6.56	66.69 ± 7.91	<i>P</i> < 0.001 ^c
Gender (male) ^b	293 (65.0)	121 (58.5)	<i>P</i> = 0.507
FEV1%pred ^a	50.40 ± 15.11	57.06 ± 21.96	<i>P</i> = 0.012 ^c
GOLD I ^b	-	35 (17.4)	<i>P</i> = 0.239
GOLD II	227 (50.3)	80 (39.8)	
GOLD III	176 (39.0)	61 (30.3)	
GOLD IV	48 (10.6)	25 (12.4)	
Smoking pack years ^a	40.00 (30.00–50.00)	37.5 (22.50–51.25)	<i>P</i> = 0.108
Baseline CAT ^a	20.23 ± 7.33	18.32 ± 7.22	<i>P</i> = 0.003 ^c
Baseline CCQ Total ^a	2.86 ± 1.17	2.12 ± 1.02	<i>P</i> < 0.001 ^c
Baseline SGRQ Total ^a	50.69 ± 17.33	42.88 ± 19.16	<i>P</i> < 0.001 ^c
Baseline mMRC ^a	2 (2–4)	1 (1–2)	<i>P</i> < 0.001 ^c
Δ CAT 12 months ^a	−0.89 ± 7.01	−0.14 ± 4.92	<i>P</i> = 0.224
Δ CCQ Total 12 months ^a	−0.16 ± 1.08	0.02 ± 0.69	<i>P</i> = 0.077
Δ SGRQ Total 12 months ^a	−3.94 ± 15.38	0.87 ± 11.55	<i>P</i> < 0.001 ^c

Abbreviations: CAT, COPD Assessment Test; CCQ, Clinical COPD Questionnaire; COPD, chronic obstructive pulmonary disease; FEV1%pred, Forced Expiratory Volume in one second % predicted; GOLD, Global initiative for Chronic Obstructive Lung Disease; mMRC, modified Medical Research Council Dyspnea Scale; *N*, number of patients; SGRQ, St. George's Respiratory Questionnaire.

^a Data expressed as mean ± standard deviation or median (interquartile range).

^b Data expressed as frequencies (% of total).

^c Significance level *P* < 0.05. All listed *P*-values were corrected for multiplicity using the Benjamini–Hochberg method.

variables, the median was used to dichotomize patients into subgroups [24,28,34,35,69]. To control for the possible impact of the baseline level, relative MCIDs were also calculated as a percentage of change from baseline [21,29,36]. Finally, 0.5SDs were determined for all patients and the respective subgroups.

Multiple linear regression modeling was applied to quantify the impact of the various factors (baseline health status, gender, age, GOLD classification, and study setting) on the MCID estimates, including the analysis of possible interaction terms.

3. Results

3.1. Baseline characteristics

This retrospective analysis included 451 patients from PR of whom 309 patients completed follow-up and 207 patients from RCP of whom 177 completed follow-up (Table 1).

3.2. Global Rating of Change

Correlations between the GRC and health status change scores ($n = 2,299$) were -0.38 (CAT), -0.45 (CCQ), and -0.52 (SGRQ). The scores on the GRC differed significantly between both study settings ($P < 0.001$) (Fig. 2).

ANOVA tests between GRC groups were all significant ($P < 0.001$).

3.3. Study setting

Mean MCID estimates were for all patients for improvement and deterioration -2.82 vs. 2.66 (CAT), -0.48 vs. 0.43 (CCQ), and -8.30 vs. 5.95 (SGRQ), respectively (Table 2). Estimates were larger during PR than RCP for improvement on the CCQ (-0.49 vs. -0.40 , $P = 0.687$) and SGRQ (-8.71 vs. -3.04 , $P = 0.020$), and for deterioration on the CAT (3.44 vs. 1.50 , $P = 0.023$), CCQ (0.51 vs. 0.30 , $P = 0.090$), and SGRQ (6.11 vs. 5.69 , $P = 0.825$) (Table 2). ROC curves demonstrated a similar pattern, although estimates for deterioration were overall smaller than for improvement (Table 3). The 0.5SD estimates were also larger during PR than RCP (CAT 3 vs. 2 points; CCQ 0.5 vs. 0.3 points; and SGRQ 6–7 vs. 4–5 points) (Table 4). ICCs of the health status change scores and percent regression to the mean ranged 0.59–0.68 (17.66–23.12%) for PR and 0.75–0.89 (5.50–13.69%) for RCP.

3.4. Baseline health status

Better and worse baseline levels were grouped according to the median (CAT 20, CCQ 2.60, and SGRQ 49.09). MCID estimates for improvement were significantly smaller and for deterioration significantly larger in patients with a lower

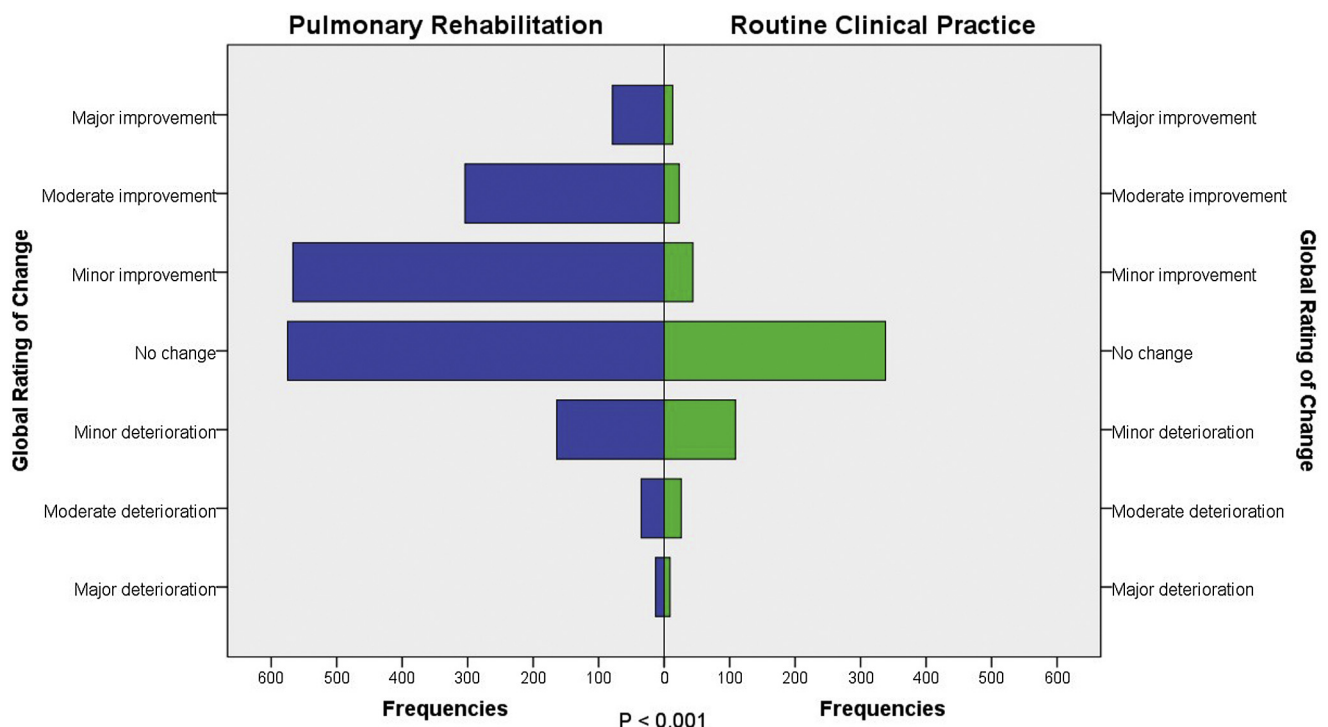


Fig. 2. Distribution of the GRC scores during PR and RCP. The blue bars (left half) represent the observed frequencies per GRC category during PR. The green bars (right half) represent the observed frequencies per GRC category during RCP. The GRC scores differed significantly between both study settings ($P < 0.001$). Abbreviations: GRC, Global Rating of Change; PR, pulmonary rehabilitation; RCP, routine clinical practice.

Table 2. Anchor-based MCID estimates for CAT, CCQ, and SGRQ using the mean change method

Patients	N	CAT Improvement	CCQ Improvement	SGRQ Improvement	N	CAT Deterioration	CCQ Deterioration	SGRQ Deterioration
All patients	611	−2.82 (−3.30 to −2.35)	−0.48 (−0.55 to −0.41)	−8.30 (−9.32 to −7.27)	272	2.66 (1.99 to 3.32)	0.43 (0.33 to 0.53)	5.95 (4.59 to 7.30)
PR patients	567	−2.81 (−3.32 to −2.31)	−0.49 (−0.57 to −0.41)	−8.71 (−9.79 to −7.63)	163	3.44 (2.48 to 4.39)	0.51 (0.37 to 0.66)	6.11 (4.16 to 8.06)
RCP patients	44	−2.91 (−4.27 to −1.55)	−0.40 (−0.58 to −0.21)	−3.04 (−5.52 to −0.57)	109	1.50 (0.70 to 2.30)	0.30 (0.19 to 0.42)	5.69 (3.95 to 7.44)
Significance	611	$P = 0.930$	$P = 0.687$	$P = 0.020^a$	272	$P = 0.023^a$	$P = 0.090$	$P = 0.825$
Better baseline health status	291 299 317	−0.67 (−1.32 to −0.03)	−0.13 (−0.23 to −0.04)	−4.66 (−6.02 to −3.29)	88 96 95	6.30 (5.07 to 7.52)	0.84 (0.68 to 0.99)	12.86 (10.63 to 15.08)
Worse baseline health status	320 312 290	−4.78 (−5.39 to −4.16)	−0.82 (−0.91 to −0.72)	−12.28 (−13.70 to −10.87)	184 175 173	0.92 (0.26 to 1.58)	0.21 (0.09 to 0.33)	2.15 (0.71 to 3.59)
Significance	611 611 607	$P < 0.001^a$	$P < 0.001^a$	$P < 0.001^a$	272 272 268	$P < 0.001^a$	$P < 0.001^a$	$P < 0.001^a$
Males	370	−2.56 (−3.17 to −1.94)	−0.38 (−0.47 to −0.29)	−8.01 (−9.30 to −6.72)	176	2.74 (1.99 to 3.49)	0.54 (0.42 to 0.67)	6.79 (4.95 to 8.63)
Females	241	−3.23 (−3.97 to −2.48)	−0.63 (−0.75 to −0.51)	−8.74 (−10.43 to −7.06)	96	2.51 (1.20 to 3.82)	0.23 (0.08 to 0.38)	4.43 (2.56 to 6.30)
Significance	611	$P = 0.544$	$P = 0.028^a$	$P = 0.655$	272	$P = 0.872$	$P = 0.011^a$	$P = 0.193$
Age low	345	−2.49 (−3.12 to −1.86)	−0.41 (−0.51 to −0.31)	−7.98 (−9.27 to −6.70)	128	3.18 (2.23 to 4.13)	0.50 (0.34 to 0.67)	6.02 (3.83 to 8.22)
Age high	266	−3.25 (−3.98 to −2.53)	−0.57 (−0.68 to −0.46)	−8.71 (−10.38 to −7.04)	144	2.19 (1.26 to 3.13)	0.37 (0.25 to 0.48)	5.88 (4.19 to 7.56)
Significance	611	$P = 0.413$	$P = 0.099$	$P = 0.635$	272	$P = 0.985$	$P = 0.370$	$P = 0.916$
GOLD I-II	310	−3.01 (−3.72 to −2.31)	−0.53 (−0.62 to −0.43)	−9.90 (−11.28 to −8.52)	137	2.35 (1.47 to 3.23)	0.44 (0.29 to 0.59)	4.86 (3.04 to 6.67)
GOLD III-IV	299	−2.63 (−3.27 to −1.98)	−0.44 (−0.55 to −0.33)	−6.64 (−8.15 to −5.12)	135	2.97 (1.97 to 3.97)	0.42 (0.30 to 0.55)	7.10 (5.07 to 9.13)
Significance	609	$P = 0.661$	$P = 0.425$	$P = 0.062$	272	$P = 0.670$	$P = 0.976$	$P = 0.224$
Co-morbidities Low RCP	32	−3.16 (−4.89 to −1.43)	−0.41 (−0.63 to −0.19)	−3.37 (−6.54 to −0.20)	64	1.20 (0.14 to 2.27)	0.30 (0.15 to 0.44)	4.42 (2.15 to 6.69)
Co-morbidities High RCP	12	−2.25 (−4.52 to −0.02)	−0.35 (−0.70 to −0.01)	−2.17 (−6.21 to +1.88)	45	1.91 (0.66 to 3.16)	0.32 (0.12 to 0.51)	7.56 (4.81 to 10.31)
Significance	44	$P = 0.778$	$P = 0.888$	$P = 0.748$	109	$P = 0.642$	$P = 0.941$	$P = 0.187$

Abbreviations: 95% CI, 95% confidence interval; CAT, COPD Assessment Test; CCQ, Clinical COPD Questionnaire; COPD, chronic obstructive pulmonary disease; GOLD, Global initiative for Chronic Obstructive Lung Disease; N, number of patients; PR, pulmonary rehabilitation; RCP, routine clinical practice; SGRQ, St. George's Respiratory Questionnaire.

Data presented as mean change scores (95% CI).

Scores marked in lighter green represent smaller MCID estimates, whereas darker green represent larger MCID estimates.

^aSignificance level $P < 0.05$ using independent t -tests. All listed P -values were corrected for multiplicity using the Benjamini–Hochberg method.

(meaning better) baseline health status during both PR and RCP ($P < 0.001$; Table 2; Fig. 3; Supplementary Table 1). Improvement thresholds for lower (meaning better) baseline compared with higher (meaning worse) baseline were for CAT −0.67 vs. −4.78, for CCQ −0.13 vs. −0.82, and for

SGRQ −4.66 vs. −12.18. Thresholds for deterioration were for CAT 6.30 vs. 0.92, for CCQ 0.84 vs. 0.21, and for SGRQ 12.86 vs. 2.15. ROC curves confirmed the pattern, although differences were less extreme between baseline severity groups (Table 3). The 0.5SD method did not show large

Table 3. Anchor-based MCID estimates for CAT, CCQ, and SGRQ using ROC curves

Patients	CAT	CCQ	SGRQ	CAT	CCQ	SGRQ
	Improvement N = 1030	Improvement N = 1030	Improvement N = 1030	Deterioration N = 356	Deterioration N = 356	Deterioration N = 356
All patients	–2.50 AUC 0.650	–0.50 AUC 0.693	–7.48 AUC 0.727	+1.50 AUC 0.671	+0.20 AUC 0.679	+4.12 AUC 0.676
PR patients	–2.50 AUC 0.644	–0.50 AUC 0.687	–7.48 AUC 0.713	+1.50 AUC 0.685	+0.25 AUC 0.661	+4.12 AUC 0.677
RCP patients	–3.50 AUC 0.650	–0.17 AUC 0.640	–3.26 AUC 0.645	+1.50 AUC 0.651	+0.10 AUC 0.723	+3.57 AUC 0.676
Better baseline health status	–1.50 AUC 0.660	–0.17 AUC 0.673	–6.14 AUC 0.746	+2.50 AUC 0.750	+0.20 AUC 0.751	+7.10 AUC 0.796
Worse baseline health status	–3.50 AUC 0.665	–0.60 AUC 0.743	–9.25 AUC 0.729	+0.50 AUC 0.706	+0.10 AUC 0.699	+0.33 AUC 0.697
Males	–2.50 AUC 0.661	–0.40 AUC 0.686	–6.49 AUC 0.731	+0.50 AUC 0.683	+0.20 AUC 0.689	+4.15 AUC 0.687
Females	–2.50 AUC 0.630	–0.55 AUC 0.697	–7.15 AUC 0.717	+0.50 AUC 0.647	+0.20 AUC 0.660	+1.97 AUC 0.655
Age Low	–1.50 AUC 0.635	–0.50 AUC 0.667	–6.47 AUC 0.710	+0.50 AUC 0.701	+0.25 AUC 0.673	+4.14 AUC 0.692
Age High	–2.50 AUC 0.672	–0.75 AUC 0.724	–7.10 AUC 0.740	+1.50 AUC 0.643	+0.10 AUC 0.686	+3.93 AUC 0.661
GOLD I-II	–2.50 AUC 0.670	–0.60 AUC 0.716	–7.52 AUC 0.765	+0.50 AUC 0.659	+0.10 AUC 0.670	+4.15 AUC 0.669
GOLD III-IV	–1.50 AUC 0.629	–0.20 AUC 0.666	–7.30 AUC 0.683	+0.50 AUC 0.681	+0.10 AUC 0.683	+3.14 AUC 0.677
Comorbidities Low	–3.50 AUC 0.624	–0.50 AUC 0.608	–4.69 AUC 0.628	+0.50 AUC 0.648	+0.20 AUC 0.736	+4.39 AUC 0.641
Comorbidities High	–1.50 AUC 0.710	–0.15 AUC 0.723	–1.78 AUC 0.684	+0.50 AUC 0.643	+0.30 AUC 0.695	+6.03 AUC 0.722

Abbreviations: AUC, area under the curve; CAT, COPD Assessment Test; CCQ, Clinical COPD Questionnaire; COPD, chronic obstructive pulmonary disease; GOLD, Global initiative for Chronic Obstructive Lung Disease; N, number of patients with important change; PR, pulmonary rehabilitation; RCP, routine clinical practice; ROC, receiver operating characteristics; SGRQ, St. George's Respiratory Questionnaire.

Data presented as MCID estimates from the ROC curves including the AUC.

Scores marked in lighter green represent smaller MCID estimates, whereas darker green represent larger MCID estimates.

differences between both groups (Table 4; Supplementary Table 2). Relative MCID estimates compared with baseline were for improvement and deterioration –7.42% and +19.25% (CAT); –11.29% and +22.65% (CCQ), and –14.31% and +18.84% (SGRQ), respectively. Their confidence intervals included for improvement –10% and for deterioration +20%. Percentage regression to the mean for low-baseline (*meaning better*) patients ranged from 31.66 to 37.55% (PR: 29.36–37.55%; RCP: 38.36–48.52%), and from 36.91 to 39.67% (PR: 38.52–40.84%; RCP: 28.80–34.20%) for high-baseline (*meaning worse*) patients.

3.5. Other variables

Compared with males, females noted larger MCID estimates for improvement (CAT –3.23 vs. –2.56, $P = 0.544$; CCQ –0.63 vs. –0.38, $P = 0.028$; SGRQ –8.74 vs. –8.01, $P = 0.655$) and smaller ones for deterioration (CAT 2.51 vs. 2.74, $P = 0.872$; CCQ 0.23 vs. 0.54, $P = 0.011$; SGRQ 4.43 vs. 6.79, $P = 0.193$) during both PR and RCP (Table 2, Supplementary Table 1). ROC curves partly confirmed this, yet MCID estimates for deterioration were smaller than for improvement (Table 3). The 0.5SD

Table 4. Distribution-based MCID estimates for CAT, CCQ, and SGRQ using the half standard deviation

Patients	CAT	CCQ	SGRQ	CAT	CCQ	SGRQ
	<i>Improved</i>	<i>Improved</i>	<i>Improved</i>	<i>Deteriorated</i>	<i>Deteriorated</i>	<i>Deteriorated</i>
All patients	−2.99	−0.46	−6.43	2.79	0.42	5.65
PR patients	−3.04	−0.47	−6.54	3.10	0.47	6.28
RCP patients	−2.24	−0.30	−4.08	2.11	0.30	4.54
Better baseline health status	−2.81	−0.41	−6.17	2.88	0.39	5.46
Worse baseline health status	−2.81	−0.44	−6.12	2.28	0.39	4.80
Males	−3.02	−0.44	−6.30	2.52	0.43	6.12
Females	−2.94	−0.48	−6.63	3.23	0.37	4.62
Age Low	−2.97	−0.47	−6.07	2.71	0.48	6.26
Age High	−3.01	−0.45	−6.89	2.84	0.36	5.07
GOLD I–II	−3.15	−0.44	−6.16	2.61	0.45	5.41
GOLD III–IV	−2.82	−0.49	−6.64	2.95	0.38	5.85
Comorbidities Low	−2.40	−0.31	−4.40	2.13	0.28	4.51
Comorbidities High	−1.79	−0.28	−3.18	2.09	0.33	4.47

Abbreviations: CAT, COPD Assessment Test; CCQ, Clinical COPD Questionnaire; COPD, chronic obstructive pulmonary disease; GOLD, Global initiative for Chronic Obstructive Lung Disease; PR, pulmonary rehabilitation; RCP, routine clinical practice; SGRQ, St. George's Respiratory Questionnaire.

Data presented as half standard deviation MCID estimates of the health status change score.

Scores marked in lighter green represent smaller MCID estimates, whereas darker green represent larger MCID estimates.

estimates were similar for gender, except for deterioration on the CAT and SGRQ (Table 4, Supplementary Table 2).

Younger patients had, compared with older patients (median 60 years), smaller MCID estimates for improvement (CAT −2.49 vs. −3.25, $P = 0.413$; CCQ −0.41 vs. −0.57, $P = 0.099$; SGRQ −7.98 vs. 8.71, $P = 0.635$), but larger estimates for deterioration (CAT 3.18 vs. 2.19, $P = 0.985$; CCQ 0.50 vs. 0.37, $P = 0.370$; SGRQ 5.88 vs. 6.02, $P = 0.916$) (Table 2). This pattern was different in RCP (Supplementary Table 1). ROC curves confirmed this pattern, except for deterioration on the CAT (Table 3). The 0.5SD estimates were consistent between both age groups (Table 4).

GOLD III–IV patients scored smaller MCID estimates for improvement (CAT −2.63 vs. −3.01, $P = 0.661$; CCQ −0.44 vs. −0.53, $P = 0.425$; and SGRQ −6.64 vs. −9.90, $P = 0.062$) and larger ones for deterioration (CAT 2.97 vs. 2.35, $P = 0.670$; SGRQ 7.10 vs. 4.86, $P = 0.224$) (Table 2). The pattern was different for improvement in RCP (Supplementary Table 1). ROC curves confirmed the pattern for improvement (Table 3). The 0.5SD estimates were consistent between both GOLD groups (Table 4).

3.6. Linear multiple regression analysis and interaction

Supplementary Table 3 demonstrates the best regression models for the MCID estimates of the CAT, CCQ, and SGRQ. Baseline health status and study setting were frequent, significant, and independent factors in most models. Various interactions were noted. In general, females had worse baseline health status; baseline health status was worse in PR; patients

in PR were younger; younger patients had worse baseline health status and GOLD classification; worse GOLD classification patients had worse baseline health status. Proportion of the variance of the MCIDs (R^2) explained in these models was between 0.205 and 0.405.

4. Discussion

4.1. Main findings

The present study demonstrated first of all that MCID estimates for improvement on the CAT, CCQ, and SGRQ were significantly three to seven times larger for COPD patients with worse baseline health status than for those with better baseline health status; however, they were much smaller for deterioration. Second, MCID estimates proved to be larger during intervention, in this case PR, compared with RCP possibly due to baseline differences. Females, elderly, COPD GOLD I and II patients, and patients with fewer comorbidities had overall larger MCID estimates for improvement and smaller ones for deterioration compared with their counter groups—although not necessarily all significant. Complex interactions between the variables were observed.

4.2. Interpretation of outcomes

4.2.1. MCID estimates for CAT, CCQ, and SGRQ

Most MCID estimates for the CAT for improvement and deterioration were between ± 1.50 and ± 3.50 , which is in

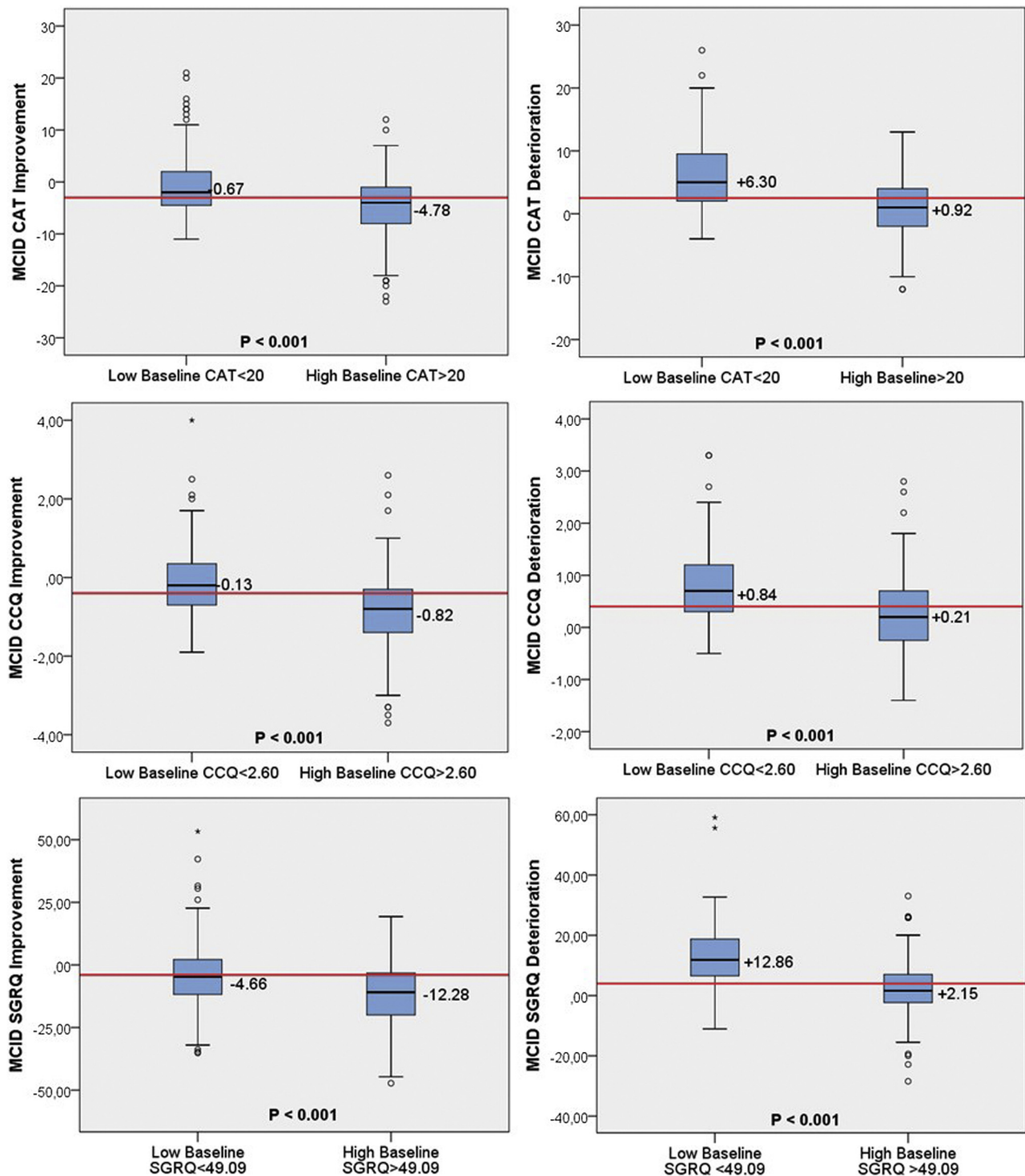


Fig. 3. Box plots of the MCID estimates for both improvement and deterioration defined by low-baseline (*meaning better*) and high-baseline (*meaning worse*) health status. Box plots of the MCID estimates for the CAT, CCQ, and SGRQ were grouped per health status baseline severity category. The left graphs represent MCID estimates for improvement, and the right half represent MCID estimates for deterioration. The red horizontal line represents the currently accepted fixed MCIDs from the literature (CAT 2 points; CCQ 0.40 points; SGRQ 4 points). Abbreviations: CAT, COPD Assessment Test; CCQ, Clinical COPD Questionnaire; COPD, chronic obstructive pulmonary disease; MCID, minimal clinically important difference; SGRQ, St. George's Respiratory Questionnaire.

accordance with the existing literature [40,43–49]. MCID estimates for the CCQ were overall in the range of ± 0.30 and ± 0.60 , which also matches previous studies [40,43,44,48–52]. SGRQ MCID estimates ranged mostly between four and nine points, which to some extent matched existing evidence [40,43,44,53–57]. An estimate of four points is generally accepted in interpreting the relevance of clinical trial outcomes. This estimate derived among others from patients treated with salmeterol in a clinical trial, but also from PR and hospital admitted patients [53–56]. The present study confirms that this four-point estimate could potentially be valid in RCP, but should be larger for interventions.

4.2.2. Baseline health status

MCIDs may be dependent on the initial baseline health state of the patient [16,17,20,21,23–39,69]. It did significantly impact our MCID estimates for COPD health status tools. Higher baseline scores (*meaning worse*) resulted in three to seven times larger MCID estimates for improvement and four to six times smaller estimates for deterioration compared with patients with a better baseline status. This means that these patients required a relative large reduction in symptoms and burden of disease before they felt better, and only little deterioration before they felt worse. Interaction between baseline health status and other variables (gender, study setting, age, and spirometry classification) was observed, potentially influencing this observed MCID pattern.

Most authors recognize that patients with worse baseline scores require more change before it is to be considered clinically relevant simply because there is more room for change [16,20,24,30,31,34–39]. Small improvements are not considered important after intervention or during routine clinical practice. On the other hand, only small progression of their severe health status would be considered a relevant deterioration. Although perhaps this may be considered a predictable outcome, no former studies in COPD health status have explored this phenomenon in MCID research. Regression to the mean in our study may explain part of the outcome. Other studies considered the use of relative MCIDs—defined as change in percentage from baseline—to solve the baseline dilemma [20,21,29,36,70]. Relative MCID estimates in the present study were approximately –10% for improvement and +20% for deterioration. This could possibly be a solution and may be applied in clinical practice to interpret individual change scores.

4.2.3. COPD disease severity

It has been hypothesized that disease severity—measured in the present study by spirometry—could impact the MCID [16,17,20,23–39]. Previous—but also present—research demonstrated that worse health status was correlated with worse lung function [9–13], although this correlation was only weak to moderate [4–7]. Worse GOLD-classified patients had also worse baseline health status in our study and were in general younger. Our study suggested that MCID

estimates were larger for improvement and smaller for deterioration in GOLD I/II patients compared with GOLD III/IV patients. This pattern is vice versa the pattern found for the impact of the baseline health status severity on the MCID. Severity of health status is thus not equivalent to COPD disease severity, as expressed in the small-to-moderate correlation between spirometry and health status. It could be argued that patients with more severe lung function would experience more exacerbations and hospital admissions, which could mean that small changes in the disease state could already be considered important [33]. Age might have interacted in the pattern observed.

4.2.4. Study setting

Setting may impact an instrument's MCID [17,20,25,38], potentially leading to larger MCIDs during intervention [71]. In the present study, MCID estimates for improvement and deterioration in COPD health status were indeed larger during intervention in PR than RCP, although not all results were significant. Patients experienced more change during intervention as a result of treatment, leading to a larger MCID estimate. This perhaps predictable result has not been demonstrated in previous COPD health status research. A systematic review by Alma et al. (2018) [40] on these MCIDs could not observe a similar pattern. In RCP, smaller changes may be noted and regarded relevant. In the present study, patients during PR were significantly younger and had worse spirometry and health status at baseline. These factors could have interacted with the different MCID estimates between both settings. The sole impact of setting on the MCID can therefore not be quantified. Furthermore, the sample size during RCP was much smaller, possibly impacting estimates too. Finally, it remains unclear, whether this finding is specific for a rehabilitation intervention or is generally true for any kind of intervention. For this question, further studies would be needed.

4.2.5. Gender, age, and comorbidities

Gender, age, and comorbidities could impact health status [8–10,13–15]. First, gender was hypothesized to impact health status and its MCIDs [23]. Men and women evaluated health status differently. Females pay more attention to dyspnea, emotions, and anxiety; and they had more comorbidities [5,72]. The present study demonstrated that females had (nonsignificant) larger MCID estimates for improvement and smaller ones for deterioration. Here, the worse baseline health status of females could interact and explain our findings.

Next, age may possibly impact the MCID too [17,23]. Younger patients experienced significantly worse health status and dyspnea compared with elderly [5,73]. However, older age has also been associated with worse health status [5,14]. In the present study, MCID estimates were larger for improvement and smaller for deterioration in elderly, although not all significant. Older patients had significantly better baseline health status and spirometry. The study in RCP included significantly more elderly than during PR. However, the other interacting patterns found above cannot

explain the impact of age on the MCID. Our findings contradict the results by Arima et al. [22] that older patients had lower/smaller MCIDs. These authors hypothesized that elderly had lower expectations, being satisfied, and thus requiring smaller MCIDs.

Finally, comorbidities could impact health status and its MCIDs [17]. Patients with COPD experience a variety of comorbidities [74]. MCIDs for improvement were larger for RCP patients with fewer comorbidities and smaller in deterioration for this group. Comorbidities contribute to the overall disease severity, which would match the pattern found for the COPD disease severity defined by spirometry. Patients with fewer comorbidities were significantly younger, which could imply that age interacted in the pattern too.

4.3. Strengths and limitations

The present study is the first study to explore the impact of various factors on MCID estimates for COPD health status tools. It used a large number of observations obtained from two settings during different follow-up periods applying both anchor- and distribution-based approaches. Although the impact of study setting, intervention, and baseline health status score was perhaps predictable, no previous studies have confirmed this in COPD research. No standard approach exists to evaluate the impact of factors on the MCID. The present study dichotomized the impact factors into subgroups. This has been applied by other authors in MCID research outside the field of COPD [24,28,34,35,38,39,69]. The current authors summed up and subsequently analyzed all health status change scores simultaneously to allow for subgroup analysis. The dependency of the change scores was only moderate. Alma et al. [44] showed that the recall period was of limited influence on the MCID, supporting the validity of combining all measurements. Finally, the difference between MCID estimates was tested and *P*-values were corrected for multiple testing.

There are, however, limitations too. Overall, the patterns observed were not all significant after correcting for multiplicity. Owing to the exploratory nature of this study, we chose to report general trends. Next, impact factors on the MCID were analyzed individually; however, interactions have been observed between the various factors. This makes simple conclusions difficult to establish; especially as the explained variance R^2 was low. Furthermore, in the anchor-based approach, the GRC was used to differentiate between important and unimportant change. Although correlations fulfilled the pre-set requirements [68], observed correlations were considered weak to moderate. Moreover, it has been argued that GRC estimates could be more related with the follow-up health status score because of a response shift, therefore not certainly representing change from baseline [75]. Next, the division between PR and RCP patients was unequal, therefore providing more weight to PR measurements. GRC scoring patterns between PR and RCP were also significantly different. Setting impacted

the MCID, possibly influencing other subgroup analyses too. Finally, the study on comorbidities was only valid during RCP, as scores were not readily available for PR.

4.4. Implications for clinical practice and future research

At the group level, regression to the mean may play a major role in clinical trials. This means that less weight will be distributed to outlying measurements, balancing out extreme scores. It could be hypothesized that this will minimize the impact of individual patient-related factors and the health status baseline score on the MCID. However, if samples have extreme baseline characteristics or unbalanced divisions, subgroup analyses with clustered MCIDs would be preferred to interpret outcomes in scientific research more precise. The specific trends found for the impact of factors on the MCIDs of COPD health status tools in the present study might be a start to develop an algorithm in evaluating individual health status changes during clinical practice using tailored MCIDs. More research would be required here to confirm our findings and explore the interactive nature of the variables.

5. Conclusions

The MCID is currently used as a nonadaptable parameter in the interpretation of COPD health status during clinical trials. However, our study demonstrated that a complex interaction of study setting, baseline health status, gender, age, spirometry classification, and comorbidities potentially impacted the MCID estimates for both improvement and deterioration on three major health status questionnaires. More accurate individual interpretation of outcomes in scientific research and clinical practice would benefit from developing and using possibly clustered or even tailored MCIDs.

Acknowledgments

The authors are grateful to the Junior Scientific Masterclass (JSM) of the University of Groningen, the Netherlands, who financially supported the research position of the first author.

Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.jclinepi.2019.07.015>.

References

- [1] Jones PW. Health status measurement in chronic obstructive pulmonary disease. *Thorax* 2001;56(11):880–7.
- [2] Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management and prevention of

- COPD (2017 Report). 2017. Available at <http://www.goldcopd.org/gold-2017-global-strategy-diagnosis-Management-prevention-copd/>. Accessed July 9, 2018.
- [3] Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;118(8):622–9.
 - [4] Westwood M, Bourbeau J, Jones PW, Cerulli A, Capkun-Niggli G, Worthy G. Relationship between FEV1 change and patient-reported outcomes in randomised trials of inhaled bronchodilators for stable COPD: a systematic review. *Respir Res* 2011;12:40.
 - [5] Tsiligianni I, Kocks J, Tzanakis N, Siafakas N, van der Molen T. Factors that influence disease-specific quality of life or health status in patients with COPD: a review and meta-analysis of Pearson correlations. *Prim Care Respir J* 2011;20(3):257–68.
 - [6] Curtis JR, Patrick DL. The assessment of health status among patients with COPD. *Eur Respir J* 2003;21(Suppl. 41):36s–45s.
 - [7] Jones P, Miravittles M, van der Molen T, Kulich K. Beyond FEV1 in COPD: a review of patient-reported outcomes and their measurement. *Int J Chron Obstruct Pulmon Dis* 2012;7:697–709.
 - [8] Yoo JY, Kim YS, Kim SS, Lee HK, Park CG, Oh EG, et al. Factors affecting the trajectory of health-related quality of life in COPD patients. *Int J Tuberc Lung Dis* 2016;20(6):738–46.
 - [9] Rubinsztajn R, Przybyłowski T, Maskey-Warzechowska M, Karwat K, Chazan R. Exacerbations of chronic obstructive pulmonary disease and quality of life of patients. *Adv Exp Med Biol* 2016;884:69–74.
 - [10] Kim JK, Lee SH, Lee BH, Lee CY, Kim do J, Min KH, et al. Factors associated with exacerbation in mild-to-moderate COPD patients. *Int J Chron Obstruct Pulmon Dis* 2016;11:1327–33.
 - [11] Hensch I, Strang S, Löfdahl CG, Ekberg-Jansson A. Health-related quality of life in a nationwide cohort of patients with COPD related to other characteristics. *Eur Clin Respir J* 2016;3:31459.
 - [12] Wacker ME, Jörres RA, Karch A, Wilke S, Heinrich J, Karrasch S, et al. Assessing health-related quality of life in COPD: comparing generic and disease-specific instruments with focus on comorbidities. *BMC Pulm Med* 2016;16(1):70.
 - [13] Corlateanu A, Botnaru V, Covantev S, Dumitru S, Siafakas N. Predicting health-related quality of life in patients with chronic obstructive pulmonary disease: the impact of age. *Respiration* 2016;92(4):229–34.
 - [14] Sundh J, Montgomery S, Hasselgren M, Kämppe M, Janson C, Ställberg B, et al. Change in health status in COPD: a seven-year follow-up cohort study. *NPJ Prim Care Respir Med* 2016;26:16073.
 - [15] Nagai K, Makita H, Suzuki M, Shimizu K, Konno S, Ito YM, et al. Differential changes in quality of life components over 5 years in chronic obstructive pulmonary disease patients. *Int J Chron Obstruct Pulmon Dis* 2015;10:745–57.
 - [16] Jones PW, Beeh KM, Chapman KR, Decramer M, Mahler DA, Wedzicha JA. Minimal clinically important differences in pharmacological trials. *Am J Respir Crit Care Med* 2014;189(3):250–5.
 - [17] Maltenfort M, Diaz-Ledezma C. Statistics in brief: minimum clinically important difference – availability of reliable estimates. *Clin Orthop Relat Res* 2017;475(4):933–46.
 - [18] Leopold SS, Porcher R. Editorial: the minimum clinically important difference – the least we can do. *Clin Orthop Relat Res* 2017;475(4):929–32.
 - [19] Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10(4):407–15.
 - [20] Angst F, Aeschlimann A, Angst J. The minimal clinically important difference raised the significance of outcome effects above the statistical level, with methodological implications for future studies. *J Clin Epidemiol* 2017;82:128–36.
 - [21] Terwee CB, Roorda LD, Dekker J, Bierma-Zeinstra SM, Peat G, Jordan KP, et al. Mind the MIC: large variation among populations and methods. *J Clin Epidemiol* 2010;63:524–34.
 - [22] Arima H, Carreon LY, Glassman SD, Yamato Y, Hasegawa T, Togawa D, et al. Age variation in the minimum clinically important difference in SRS-22r after surgical treatment for adult spinal deformity – a single institution analysis in Japan. *J Orthop Sci* 2018;23(1):20–5.
 - [23] Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain* 2008;9(2):105–21.
 - [24] Wang YC, Hart DL, Stratford PW, Mioduski JE. Baseline dependency of minimal clinically important improvement. *Phys Ther* 2011;91(5):675–88.
 - [25] Katz NP, Paillard FC, Ekman E. Determining the clinical importance of treatment benefits for interventions for painful orthopedic conditions. *J Orthop Surg Res* 2015;10:24.
 - [26] Stucki G, Daltroy L, Katz JN, Johannesson M, Liang MH. Interpretation of change scores in ordinal clinical scale and health status measures: the whole may not equal the sum of the parts. *J Clin Epidemiol* 1996;49(7):711–7.
 - [27] Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. *Curr Opin Rheumatol* 2002;14(2):109–14.
 - [28] Baker DW, Hays RD, Brook RH. Understanding changes in health status: is the floor phenomenon Merely the last step of the staircase? *Med Care* 1997;35(1):1–15.
 - [29] 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(5):541–6.
 - [30] Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol* 2003;56(5):395–407.
 - [31] Hays RD, Woolley JM. The concept of clinically meaningful difference in health-related quality of life research. *Pharmacoeconomics* 2000;18(5):419–23.
 - [32] Engel L, Beaton DE, Touma Z. Minimal clinically important difference: a review of outcome measure score interpretation. *Rheum Dis Clin North Am* 2018;44(2):177–88.
 - [33] Jones PW, Rennard S, Tabberer M, Riley JH, Vahdati-Bolouri M, Barnes NC. Interpreting patient-reported outcomes from clinical trials in COPD: a discussion. *Int J Chron Obstruct Pulmon Dis* 2016;11:3069–78.
 - [34] Stratford PW, Binkley JM, Riddle DL, Guyatt GH. Sensitivity to change of the roland-Morris back pain questionnaire: part 1. *Phys Ther* 1998;78(11):1186–96.
 - [35] Schuller W, Ostelo RW, Janssen R, de Vet HC. The Influence of study population and definition of improvement on the smallest detectable change and the minimal important change of the neck disability index. *Health Qual Life Outcomes* 2014;12:53.
 - [36] Farrar JT, Young JPI, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured in an 11-point numerical pain rating scale. *Pain* 2001;94(2):149–58.
 - [37] de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health Qual Life Outcomes* 2006;4:54.
 - [38] de Vet HC, Fomani M, Scholten MA, Jacobs WC, Stiggelbout AM, Knol DL, et al. Minimally important change values of a measurement instrument depend more on baseline values than on the type of intervention. *J Clin Epidemiol* 2015;68(5):518–24.
 - [39] de Vet HC, Ostelo RW, Terwee CB, van der Roer N, Knol DL, Beckerman H, et al. Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. *Qual Life Res* 2007;16(1):131–42.
 - [40] Alma H, de Jong C, Tsiligianni I, Sanderma R, Kocks J, van der Molen T. Clinically relevant differences in COPD health status: systematic review and triangulation. *Eur Resp J* 2018;52(3):1800412.
 - [41] Cave AJ, Atkinson L, Tsiligianni IG, Kaplan AG. Assessment of COPD wellness tools for use in primary care: an IPCRG initiative. *Int J Chron Obstruct Pulmon Dis* 2012;7:447–56.

- [42] Kocks JWH, Seys SF, van Duin TS, Diamant Z, Tsiligianni IG. Assessing patient-reported outcomes in asthma and COPD patients: which can be recommended in clinical practice? *Curr Opin Pulm Med* 2018;24(1):18–23.
- [43] Alma H, de Jong C, Jelusic D, Wittmann M, Schuler M, Flokstra-de Blok B, et al. Health status instruments for patients with COPD in pulmonary rehabilitation: defining a minimal clinically important difference. *NPJ Prim Care Respir Med* 2016;26:16041.
- [44] Alma H, de Jong C, Jelusic D, Wittmann M, Schuler M, Kollen BJ, et al. Assessing health status over time: impact of recall period and anchor question on the minimal clinically important difference of COPD health status tools. *Health Qual Life Outcomes* 2018;16(1):130.
- [45] Kon SS, Canavan JL, Jones SE, Nolan CM, Clark AL, Dickson MJ, et al. Minimum clinically important difference for the COPD Assessment Test: a prospective analysis. *Lancet Respir Med* 2014;2(3):195–203.
- [46] Dodd JW, Hogg L, Nolan J, Jefford H, Grant A, Lord VM, et al. The COPD assessment test (CAT): response to pulmonary rehabilitation. A multicenter, prospective study. *Thorax* 2011;66(5):425–9.
- [47] Jones PW, Harding G, Wiklund I, Berry P, Tabberer M, Yu R, et al. Tests of the responsiveness of the COPD assessment test following acute exacerbation and pulmonary rehabilitation. *Chest* 2012;142(1):134–40.
- [48] Smid DE, Franssen FM, Houben-Wilke S, Vanfleteren LE, Janssen DJ, Wouters EF, et al. Responsiveness and MCID estimates for CAT, CCQ and HADS in patients with COPD undergoing pulmonary rehabilitation: a prospective analysis. *J Am Med Dir Assoc* 2016;18(1):53–8.
- [49] Tsiligianni IG, van der Molen T, Moraitaki D, Lopez I, Kocks JW, Karagiannis K, 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.
- [50] Canavan JL, Dilaver D, Clark AL, Jones SE, Nolan CM, Kon SS, et al. Clinical COPD Questionnaire in patients with chronic respiratory disease. *Respirology* 2014;19(7):1006–12.
- [51] Kocks JW, Tuinenga MG, Uil SM, van den Berg JW, Stahl E, van der Molen T. Health status measurement in COPD: the minimal clinically important difference of the clinical COPD questionnaire. *Respir Res* 2006;7:62.
- [52] Kon SS, Dilaver D, Mittal M, Nolan CM, Clark AL, Canavan JL, et al. The Clinical COPD Questionnaire: response to pulmonary rehabilitation and minimal clinically important difference. *Thorax* 2014;69(9):793–8.
- [53] Osman IM, Godden DJ, Friend JA, Legge JS, Douglas JG. Quality of life and hospital re-admission in patients with chronic obstructive pulmonary disease. *Thorax* 1997;52(1):67–71.
- [54] Jones PW, Bosch TK. Quality of life changes in COPD patients treated with Salmeterol. *Am J Respir Crit Care Med* 1997;155(4):1283–9.
- [55] Jones PW. St. George's respiratory questionnaire: MCID. *J Chronic Obstructive Pulm Dis* 2005;2(1):75–9.
- [56] Schunemann HJ, Griffith L, Jaeschke R, Goldstein R, Stubbings 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(12):1170–6.
- [57] Welling JB, Hartman JE, Ten Hacken NH, Klooster K, Slebos DJ. The minimal important difference for the St George's Respiratory Questionnaire in patients with severe COPD. *Eur Respir J* 2015;46(6):1598–604.
- [58] Schultz K, Jelusic D, Wittmann M, Krämer B, Huber V, Fuchs S, et al. Inspiratory Muscle training does not improve clinical outcomes in 3-week COPD rehabilitation: results from a randomised controlled trial. *Eur Respir J* 2018;51(1):1702000.
- [59] Alma HJ, de Jong C, Jelusic D, Wittmann M, Schuler M, Sanderman R, et al. Thresholds for clinically important deterioration versus improvement in COPD health status: results from a randomised controlled trial in pulmonary rehabilitation and an observational study during routine clinical practice. *BMJ Open* 2019;9(6):e025776.
- [60] 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(3):648–54.
- [61] Van der Molen T, Willemse BW, Schokker S, ten Hacken NH, Postma DS, Juniper EF. Development, validity and responsiveness of the clinical COPD questionnaire. *Health Qual Life Outcomes* 2003;1:13.
- [62] Jones PW, Quirk FH, Baveystock CM. The St George's respiratory questionnaire. *Respir Med* 1991;85(Suppl B):25–31.
- [63] 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(1):81–7.
- [64] de Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: a practical guide. Cambridge: Cambridge University Press; 2011.
- [65] Norman GR, Sloan JA, Wywich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003;41(5):582–92.
- [66] Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60(1):34–42.
- [67] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 1995;57:289–300.
- [68] 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(2):102–9.
- [69] Rouquette A, Blanchin M, Sébille V, Guillemin F, Coté SM, Falissard B, et al. The minimal clinically important difference determined using item response theory models: an attempt to solve the issue of the association with baseline score. *J Clin Epidemiol* 2014;67(4):433–40.
- [70] Zhang Y, Zhang S, Thabane L, Furukawa TA, Johnston BC, Guyatt GH. Although not consistently superior, the absolute approach to framing the minimally important difference has advantages over the relative approach. *J Clin Epidemiol* 2015;68(8):888–94.
- [71] Yuksel S, Ayhan S, Nabiye V, Domingo-Sabat M, Vila-Casademunt A, Obeid I, et al. Minimum clinically important difference of the health-related quality of life scales in adult spinal deformity calculated by latent class Analysis: is it appropriate to use the same values for surgical and non-surgical patients? *Spine J* 2019;19(1):71–8.
- [72] Nicolini A, Barbagelata E, Tagliabue E, Colombo D, Monacelli F, Braido F. Gender differences in chronic obstructive pulmonary diseases: a narrative review. *Panminerva Med* 2018;60(4):192–9.
- [73] Martinez CH, Diaz AA, Parulekar AD, Rennard SI, Kanner RE, Hansel NN, et al. Age-related differences in health-related quality of life in COPD. *Chest* 2016;149(4):927–35.
- [74] Corlateanu A, Covantev S, Mathioudakis AG, Botnaru V, Siafakas N. Prevalence and burden of comorbidities in chronic obstructive pulmonary disease. *Respir Invest* 2016;54(6):387–96.
- [75] Ousmen A, Conroy T, Guillemin F, Velten M, Jolly D, Mercier M, et al. Impact of the occurrence of a response shift on the determination of the minimal important difference in a health-related quality of life score over time. *Health Qual Life Outcomes* 2016;14(1):167.