A practice test and selection of a core set of outcome-based quality indicators in Dutch primary care physical therapy for patients with COPD: a cohort study

Arie C. Verburg [1], Simone A. van Dulmen[1], Henri Kiers[2,3], Maria W.G. Nijhuis-van der Sanden[1] and Philip J. van der Wees[1]

[1]Radboud University Medical Center, Radboud Institute for Health Sciences, IQ Healthcare, Nijmegen, the Netherlands. [2]Utrecht University of Applied Sciences, Institute of Human Movement Studies, Utrecht, the Netherlands. [3]Association for Quality in Physical Therapy (SKF), Zwolle, the Netherlands.

Corresponding author: Arie C. Verburg (koen.verburg@radboudumc.nl)

Shareable abstract (@ERSpublications)
The major finding of this study is that all participants in the focus groups accepted the quality indicators as a quality improvement tool based on their perceived added value and selected a core set of seven outcome-based quality indicators for COPD. https://bit.ly/3NJuQtq

Cite this article as: Verburg AC, van Dulmen SA, Kiers H, *et al*. A practice test and selection of a core set of outcome-based quality indicators in Dutch primary care physical therapy for patients with COPD: a cohort study. *ERJ Open Res* 2022; 8: 00008-2022 [DOI: 10.1183/23120541.00008-2022].

Received: 11 Jan 2022
Accepted: 6 June 2022

Abstract

Aim To estimate the comparability and discriminability of outcome-based quality indicators by performing a practice test in Dutch physical therapy primary care, and to select a core set of outcome-based quality indicators that are well accepted by physical therapists based on their perceived added value as a quality improvement tool.

Methods First, a list of potential quality indicators was defined, followed by determination of the comparability (case-mix adjusted multilevel analysis) and discriminability (intraclass correlation coefficient (ICC)). Second, focus group meetings were conducted with stakeholders (physical therapists and senior researchers) to select a core set of quality indicators.

Results Overall, 229 physical therapists from 137 practices provided 2651 treatment episodes. Comparability: in 10 of the 11 case-mix adjusted models, the ICC increased compared with the intercept-only model. Discriminability: the ICC ranged between 0.01 and 0.34, with five of the 11 ICCs being >0.10. The majority of physical therapists in each focus group preferred the inclusion of seven quality indicators in the core set, including three process and four outcome indicators based upon the 6-min walk test (6MWT), the Clinical COPD Questionnaire (CCQ), and the determination of quadriceps strength using a hand-held dynamometer.

Conclusion This is the first study to describe the comparability and discriminability of the outcome-based quality indicators selected for patients with COPD treated in primary care physical therapy practices. Future research should focus on increasing data collection in daily practice and on the development of tangible methods to use as the core set of a quality improvement tool.

Background

The routine use of outcome measures can play an important role in improving healthcare quality [1]; for example, they can enable comparison of the performances of providers to stimulate improvement initiatives [2]. A fundamental prerequisite of the use of outcome measures is the collection, aggregation and comprehensive presentation of data that is understandable [1]. Using quality indicators may stimulate the routine data collection of patient-reported outcome measures (PROMs) by healthcare providers. Quality indicators can be used on an aggregated level to show changes in clinical practice over time [1, 3, 4].

In a previous study, we developed a standard set of outcome domains and associated measures, including PROMs and physical performance measures, for patients with COPD in primary care physical therapy

practice [4]. However, it is still unclear which quality indicators can be selected from the standard set and which quality indicators have perceived added value as quality improvement tools for such patients.

In this study we focused on outcome-based quality indicators chosen from the standard set of PROMs and physical performance measures for patients with COPD [5]. PROMs are often combined with other clinician-assessed, impairment-based or physical performance-based measures, such as the 6-min walk test (6MWT), to provide a more complete interpretation of patient outcomes [6].

Currently, most strategies for the development of quality indicators are based on an evidence-based consensus between stakeholders in procedures, such as the RAND/UCLA Delphi procedure [7–9]. This is true for recommendations for clinical practice guidelines too, such as the recently published Dutch clinical practice guideline (CPG) for primary care physical therapists treating patients with COPD [10], and can provide an important basis for the development of quality indicators [11]. In addition, a practice test, including the collection of real-world data prior to selection, is an essential step for evaluation of the comparability, discriminability and feasibility of potential quality indicators in daily practice [9, 12]. A practice test can support the usefulness and feasibility of quality indicators in daily practice and gain insight into the psychometric properties of outcome-based quality indicators [3, 11, 13]. Although to our knowledge, no specific definition of a practice test has been reported in previous research, there are several examples of using a practice test in the development of quality indicators [11, 13, 14]. Such an example is the study of by Meerhoff *et al*. [13] in which a practice test was conducted to explore the reliability, validity and discriminability of patient-reported outcomes for the development of quality indicators in patients with nonspecific low back pain. We defined comparability as the extent to which the quality indicator is comparable between practices, and discriminability as the extent to which the quality indicator is able to discriminate between practices.

Here, we develop outcome-based quality indicators for patients with COPD in physical therapy primary care. The aims of this study are therefore: 1) to estimate the comparability and discriminability of outcome-based quality indicators included in a previously selected standard set of measures; 2) to select a core set of outcome-based quality indicators that is well accepted by physical therapists based on the perceived added value of this core set as a quality improvement tool.

Methods

Design

This mixed methods study used a sequential explanatory design, taking a previously selected standard set of outcome domains and measures as the basis for defining and selecting a core set of quality indicators. The standard set was developed in two consecutive steps between February 2018 and April 2020 [4], and was registered on the Core Outcome Measures in Effectiveness Trials (COMET) website [15]. In phase 1 of the present study, potential quality indicators were defined, and we estimated their comparability and discriminability with prospectively collected cohort data between February 2018 and December 2019. To enhance the comparability, we adjusted for differences in patient characteristics using a case-mix correction. Furthermore, we calculated whether the quality indicator was able to discriminate the outcomes of patients between practices and could therefore be used as an instrument for quality improvement. In phase 2, we explored the perceived added value of the indicators in focus group meetings with physical therapists. We then actively involved participants in the selection of a core set of quality indicators.

Setting

A total of 229 Dutch physical therapists working in 137 primary care practices collected the treatment outcomes of patients with COPD. All participants in the project were recruited *via* stakeholder organisations in Dutch primary physical therapy care. Participating physical therapists were instructed to treat their patients according to Dutch clinical guideline recommendations for patients with COPD [16]. We only measured outcomes of the treatment; the physical therapists individually decided which treatment was needed for their patients. All procedures were conducted according to the Declaration of Helsinki and approved by the Medical Ethical Committee of Radboud University Medical Center (registration # 2019-5455). The STROBE checklist was used to report the current study [17]. Furthermore, a framework with tools to support the selection and implementation of PROMs was used as guidance for conduction of this study [3].

Data collection

Data on the treatment outcomes were anonymously collected through electronic health records (EHRs) *via* three databases: the national data registry (LDK) of the Association for Quality in Physical Therapy (SKF), the national data registry (LDF) of the Royal Dutch Society for Physical Therapy (KNGF), and the

database of Spot On Medics (SOM), which is one of the EHR vendors. The EHRs uploaded to the national registries only contain anonymised data. Furthermore, to ensure the uniformity of the provided data, all data in the registries were collected based on predefined technical specifications [18]. Informed consent was obtained and registered in the EHR from all participating patients included in the current study.

Outcome domains and measures

The outcome domains in the standard set were based on the consensus between stakeholders (patients, physical therapists, policy-makers, researchers and health insurers) [4], and on the recommendations in the Dutch CPG for the physical therapy treatment of patients with COPD [16]. After development of the standard set, the KNGF published an update for this CPG [10], in which the suggested outcome domains and associated measures to evaluate physical therapy treatment are in line with the outcome domains from the developed standard set.

The standard set of outcome measures consisted of three mandatory measures for the total population, two conditional measures that depended on the treatment goal, and two exploratory measures that were used as a pilot in a small subgroup. In the current study, only the mandatory and conditional measures were used for the development of quality indicators, because the exploratory measure was only used in a small subgroup of practices. The three mandatory measures for all patients with COPD were the 6MWT to measure physical capacity, the Clinical COPD Questionnaire (CCQ) to measure health-related quality of life, and the Global Perceived Effect – Dutch Version (GPE-DV) to measure the perceived effect. The two conditional measures were the hand-held dynamometer (with Microfet™) to measure quadriceps strength, and the Medical Research Council (MRC) Dyspnoea Scale to measure dyspnoea. All measures were completed pre- and post-treatment to monitor the changes in outcomes over time, except the GPE-DV, which was only measured after the treatment. For a description of each measure and the measurement protocol, see Appendix A. All physical therapists followed a specific protocol to standardise the testing procedure [19].

Inclusion and exclusion criteria

All patients with COPD (GOLD I–IV), as diagnosed by a medical doctor, who received physical therapy in one of the participating primary care practices between February 2018 and December 2019, were included. Participating physical therapy practices were instructed to collect, as a minimum, all mandatory and conditional measures from the standard set as presented in the data collection, according to the measurement protocol described in Appendix A. Based on a rule of thumb, a minimum of 30 patients should preferably be included for each practice to allow a valid comparison [20, 21]. However, it was expected that this inclusion requirement could not be reached due to the short inclusion period, and the fact that routine data collection in primary physical therapy care for patients with COPD is relatively new. We therefore used a lower threshold, and physical therapy practices were excluded from the analysis for a specific quality indicator if they included fewer than 10 patients with COPD.

Phase 1: defining quality indicators and estimating their comparability and discriminability

Defining potential quality indicators

We used national and international standards to define potential quality indicators [2, 5, 22, 23]. Quality indicators can be described using mean values and relative differences, or quantified and expressed as a proportion in which the numerator describes the number of "correct" scores and the denominator is the number of persons for which the quality indicator is applicable [3] (see table 1 for an example of a quality indicator for physical capacity measured with the 6MWT).

TABLE 1 Example of a quality indicator monitoring the repeated measurement of the 6MWT

Definition	The proportion of patients with COPD who underwent physical therapy treatment and who completed the 6MWT pre- and post-treatment to evaluate physical capacity
Rationale	Improvement of physical capacity is an important goal in physical therapy treatment for patients with COPD. Physical capacity is measured with the 6MWT
Numerator	The number of patients who underwent physical therapy treatment and who completed the 6MWT pre- and post-treatment
Denominator	All patients who underwent physical therapy treatment
Specification	Physical capacity is measured in all patients using the 6MWT, a physical performance test where the patients walk for 6 minutes in a comfortable way
Type of indicator	Process

For each of the five measures in the standard set, we defined four types of quality indicators: 1) by monitoring the process, *i.e.* whether the outcome was actually measured pre- and post-treatment; 2) by using mean end scores of the outcome, reflecting patient functioning at the end of treatment; 3) by using the mean pre- to post-treatment change in the outcome score, reflecting improvement or decline in the outcome; and 4) by using the minimal clinically important difference (MCID) of the outcome, *i.e.* the proportion of patients who experienced clinically relevant improvements, stabilisations or deteriorations (see Box 1 for an example). The change score and the MCID were not defined for the GPE-DV, as this measure was only completed after the treatment. No MCID was defined for the MRC Dyspnoea Scale, because a MCID has not yet been established for this measure [24]. Hence, in total, we defined 17 potential quality indicators (see Appendix B for an extensive description of each of these).

Estimating the comparability and discriminability of the quality indicators

For each measure, at least 30 physical therapy practices needed to be included; this was based on a rule of thumb in multilevel analysis for general calculation [20, 21]. Descriptive statistics were used to determine whether the thresholds for the completeness of the measures were met to estimate the indicator scores. For the analysis, we used measures collected at the beginning and/or end of the treatment. When the treatment episode had not ended, we used the last provided data. The treatment episode is a unique episode of a patient being treated by a physical therapist.

Comparability

In a linear and logistic multilevel analysis, patients were clustered within physical therapy practices to compare the outcomes of the quality indicators between practices [25]. The quality indicators were adjusted for patient characteristics that influence the outcome but were not under the control of the physical therapist or physical therapy practice (so-called explanatory variables) [23]. Explanatory variables, such as age, gender and baseline scores for each measure, were used for the adjustment of the multilevel analyses. In the analysis, we started with an intercept-only model that estimates only the intercept and the random variation around the intercept. The inclusion of age [26, 27], gender [26] and the baseline score [28, 29] of each measure for the adjustment of a multilevel analysis is common in the field of quality indicator development and the comparison of provider performance [2, 25–30]. Next, all explanatory variables were added to the adjusted model, and the influence of the explanatory variables was evaluated by the amount of the random intercept variance that was explained [25].

For each physical therapy practice in the case-mix adjusted multilevel analysis, the mean scores were estimated with a 95% CI.

Box 1: Potential process and outcome quality indicators at the physical therapist or practice level

a) **Process indicator: proportion of repeated measures**
The proportion of patients with COPD who underwent physical therapy treatment in which a pre- and post-measurement was used.
Example 1: In 60% of patients, physical capacity was measured pre- and post-treatment with the 6MWT
b) **Outcome indicator: mean end scores**
The mean end score (with 95% CI) of patients with COPD after a physical therapy treatment.
Example 2: The mean end score of the health-related quality of life of patients with COPD measured with the CCQ is 2.2 points (±0.9 points)
c) **Outcome indicator: mean change scores**
The mean change score (with 95% CI) of patients with COPD between the pre- and post-physical therapy stages.
Example 3: The mean change score in the symptoms of dyspnoea in patients with COPD measured with the MRC is 2.5 points of improvement (±1.0 points).
d) **Outcome indicator: MCID**
The proportion (with 95% CI) of patients with COPD who experienced a MCID improvement between the pre- and post-treatment stages.
Example 4: In 70% (±7%) of patients, a clinically relevant change in quadriceps strength was reported after treatment, as measured with the HHD.
Abbreviations: 6MWT: 6-min walk test; CCQ: Clinical COPD Questionnaire; GPE-DV: Global Perceived Effect – Dutch Version; HHD: hand-held dynamometer; MCID: minimal clinically important difference; MRC: Medical Research Council Dyspnoea Scale.

Discriminability

To estimate the variation in the outcomes between physical therapy practices, the intraclass correlation coefficient (ICC) was calculated. The ICC for physical therapy practices was defined by dividing the variance between practices by the summation of the variance between and the variance within physical therapy practices [25]. In multilevel analyses, most ICCs are between 0.05 and 0.20, and ICCs >0.10 can be interpreted as adequate, indicating that the quality indicator is able to discriminate outcomes between physical therapists or practices [25, 31]. The ICC was also used to compare the intercept-only model with the adjusted model containing the explanatory variables (case-mix).

Visual representation of indicator scores

Caterpillar plots are used to present the mean outcomes for each defined quality indicator of each physical therapy practice in one graph, as they are found to be user-friendly and easy to interpret [2, 29, 31, 32]. We used relative norms by presenting the plots in three colours: blue (95% CI significantly lower than average), purple (no significant 95% CI difference from average) and green (95% CI significantly higher than average). The plots were used to present the outcomes of the cohort data to the participants in phase 2 of the study.

Phase 2: selecting a core set of quality indicators

Semi-structured focus group interviews were conducted by purposefully selected participating physical therapists who collected data in phase 1. We also organised one focus group meeting with Dutch senior physical therapists and senior researchers who were members of the development group of the revised Dutch physical therapy guideline for COPD. The senior researchers were asked to comment on the set of quality indicators from a scientific perspective. The senior physical therapists and researchers had at least 10 years of experience in the treatment of and/or research into patients with COPD.

We aimed to conduct four focus group meetings with 6–10 members in every meeting. The primary goal of the focus groups was to reflect on the added value of using the presented indicators in daily practice for quality improvement, and most importantly, to select a core set of quality indicators from the 17 potential indicators described. The focus group meetings were audio recorded and summarised by researcher AV; the summaries of the different focus groups were discussed and interpreted in several meetings with researchers AV (physical therapist and PhD student), SvD (physical therapist and senior researcher) and PvdW (physical therapist and professor of allied health sciences). The identities of the physical therapists were considered confidential, and so the answers given by the physical therapists during the interviews and in the survey were processed anonymously. The focus groups were part of the process of reaching consensus on the selection of the core set.

The research members AV, SvD, PvdW, HK (physical therapist and senior researcher), RN (physical therapist and professor of allied health sciences) were trained and had experience in conducting qualitative research.

In each focus group meeting, we presented each potential quality indicator in a caterpillar plot, with scores at the levels of the physical therapist and the physical therapy practice, and compared them with the scores for the other practices. The participants interpreted the comparability and discriminability of the potential quality indicators. Finally, we asked participants to select their preferred quality indicators for the core set from the potential quality indicators as described in Box 1. During each meeting, the chairman (AV) summed up all the preferred quality indicators and asked the group whether they accepted or declined the proposed core set. A consensus was reached if >80% of the participants accepted the selection of each quality indicator in the core set.

Patient and public involvement

For the development of this standard set we interviewed patients about their perspectives on the selection of patient outcomes [4, 15]. Furthermore, during the conduct of this study a steering committee with representatives from important stakeholders, including the association for patients with COPD Netherlands Patients Federation, advised during the selection process. During the meetings, we discussed the views and perspectives of stakeholders regarding the value and implementation of outcome-based quality indicators for Dutch physical therapy.

Results

Phase 1: estimating the comparability and discriminability of the quality indicators

Descriptive statistics

Table 2 shows descriptive statistics of the included treatment episodes and the number of physical therapists and physical therapy practices who provided the data. The treatment episode is a unique episode

TABLE 2 Descriptive statistics of the included patients and the number of participating physical therapists and physical therapy practices

Number of treatment episodes in the dataset	4651
Female patients	2440 (52.5%)
Age, years	67.9 (9.4)
Treatment sessions	49.2 (58.2)
Episode duration, weeks	46.6 (50.3)
Physical therapists who provided the data	229
Physical therapy practices that provided the data	137

Data are presented as means (SD) or numbers and percentages of the total population.

of a patient being treated by a physical therapist. The current national data registries cannot detect recurrences of patients over time due to privacy regulations; therefore, the number of unique patients may be lower. Overall, 229 physical therapists from 137 practices provided 4651 treatment episodes of patients with COPD.

Table 3 presents the characteristics and unadjusted outcomes of patients with COPD on each measure of the standard set. The number of patients with end scores differed between the measures. Each measure reached the threshold of at least 30 included physical therapy practices that provided ⩾10 cases, except for the HHD, for which only 10 physical therapy practices provided ⩾10 cases and therefore no ICC was calculated (see table 3 for the number of practices and provided cases that were included in the multilevel analysis).

Comparability

Table 4 presents the ICC calculations of the intercept-only models and the models adjusted with the explanatory variables. For the process measures, similar to the HHD, no ICC was calculated due to insufficient data. In total, 11 models could be estimated for the outcome indicators based on scores of the 6MWT (four indicators), CCQ (four indicators), GPE (one indicator for the end score) and MRC (two indicators; MCID could not be calculated). In 10 of the 11 case-mix adjusted models, the ICC increased

TABLE 3 Descriptive characteristics and unadjusted outcomes of patients with COPD for each measure of the standard set

	6MWT	CCQ	GPE-DV	HHD	MRC
Female patients	1344 (51.1%)	1786 (52.1%)	636 (50.6%)	223 (51.3%)	1237 (52.7%)
Age, years	67.8 (9.2)	68.1 (9.4)	68.2 (9.4)	68.0 (9.2)	68.2 (9.3)
Treatment episodes with baseline scores	2628	3427	N.A.	435	2348
Treatment episodes with end scores	1822	2408	1256	218	1385
Range of scores on each measure	6–780 m	0–6 points	1–7 points	131–542 Nm	1–5 points
Baseline scores[+]	370.8 m (126.3 m)	2.4 points (0.9)	N.A.	284.2 Nm (96.4 Nm)	3.0 points (1.1)
End scores	373.5 m (130.3 m)	2.2 points (0.9)	3.5 points (1.1)	298.4 Nm (95.5)	3.0 points (1.1)
Change (T_{end}–T_0)	2.7 (86.8)	−0.1 (0.8)	N.A.[#]	8.4 (50.9)	0.2 (1.3)
MCID improvement	533 (28.7%)[§]	818 (34.0%)[f]	N.A.[#]	99 (45.4%)[##]	N.A.[¶]
MCID stabilisation	807 (44.3%)[§]	1052 (43.7%)[f]	N.A.[#]	43 (19.9%)[##]	N.A.[¶]
MCID deterioration	490 (26.9%)[§]	537 (22.3%)[f]	N.A.[#]	76 (34.7%)[##]	N.A.[¶]
Physical therapists who provided data	145 (63.3%)	202 (88.2%)	117 (51.0%)	46 (20.0%)	168 (46.7%)
Practices that provided data	86 (62.8%)	126 (92.0%)	72 (52.6%)	28 (20.4%)	107 (78.1%)
Practices that provided ⩾10 cases	44 (19.2%)	61 (26.6%)	35 (15.3%)	10 (4.4%)	43 (18.8%)
Patients included in the multilevel analysis	1679 (36.0%)	2201 (47.3%)	1110 (23.8%)	160 (3.4%)	1226 (26.4%)

Data are presented as means (SD) or numbers and percentages of patients with baseline measures.6MWT: 6-min walk test; CCQ: Clinical COPD Questionnaire; GPE-DV: Global Perceived Effect – Dutch Version; HHD: hand-held dynamometer; MRC: Medical Research Council Dyspnoea Scale; MCID: minimal clinically important difference; N.A: not applicable. [#]GPE-DV was only analysed at the end of the treatment. [¶]The MCID for the MRC is yet to be established [24]. [+]For treatment episodes with end scores. [§]For the multilevel analysis of the 6MWT, we used an MCID of ± ⩾30 m [33]. [f]For the multilevel analysis of the CCQ, we used an MCID of ± ⩾0.4 points [34]. [##]For the multilevel analysis of the HHD, we used an MCID of ± ⩾7.5 Nm [35].

TABLE 4 ICCs for the intercept-only model and adjusted model for the change, end and MCID scores for each measure of the total population in practices that provided ⩾10 patients

	Intercept-only model	Adjusted model
6MWT end score	0.08	0.17#
6MWT change score	0.00	0.01#
6MWT MCID improvement¶	0.03	0.04#
6MWT MCID deterioration¶	0.06	0.06
CCQ end score	0.11	0.20#
CCQ change score	0.06	0.09#
CCQ MCID improvement+	0.05	0.07#
CCQ MCID deterioration+	0.03	0.05#
GPE-DV end score	0.14	0.15#
MRC end score	0.08	0.12#
MRC change score	0.23	0.34#

6MWT: 6-min walk test; CCQ: Clinical COPD Questionnaire; GPE-DV: Global Perceived Effect – Dutch Version; HHD: hand-held dynamometer; ICC: intraclass correlation coefficient; MCID: minimal clinically important difference; MRC: Medical Research Council Dyspnoea Scale; N.A: not applicable. #Increase in the ICC compared with the intercept-only model following the adjustment for the case-mix variables age, gender and baseline score of the measure. ¶For the multilevel analysis of the 6MWT, we used an MCID of ± ⩾30 m [33]. +For the multilevel analysis of the CCQ, we used an MCID of ± ⩾0.4 points [34].

compared with the intercept-only model, thus improving the comparability between practices, *i.e.*, the random intercept variance of physical therapy practices increased in the adjusted models.

Discriminability

Five of the 11 case-mix adjusted ICCs were >0.10, ranging between 0.12 and 0.32, which can be interpreted as an adequate discriminability. All adjusted models were used for the visual representation of the quality indicators in the focus group interviews. Also, the quality indicators presenting the outcomes of the HHD, for which no multilevel analysis was conducted, were presented in the focus group interviews. All defined potential quality indicators are presented as caterpillar plots (fig. 2a–c). Each graph shows that a wide range of differences in outcomes exist between physical therapy practices.

Phase 2: selecting a core set of quality indicators

In total, four focus group interviews were conducted with 20 (out of 22 invited) physical therapists and three (out of five invited) senior researchers. The nonacceptance of invited participants was due to the date and time of the focus groups, which did not fit with the agenda of the potential participants. The mean duration of the focus groups was 90 min (range 80–95). Nine were female, the mean age of the participants was 39 years (range 23–60 years), and they had an average work experience of 14 years (range 1–35 years). In total, 16 of the 20 participating physical therapists also provided data for the practice test (see Appendix C for an overview of the characteristics of the participants). Almost all the participants expressed that the presented quality indicators were user-friendly and had value for quality improvement in daily practice, but several issues surrounding the presented quality indicators were also discussed.

Using patient profiles for the comparison of patient outcomes

The participants mentioned that, in future research, it would be helpful to stratify patients based on the Dutch model, a profiling system to enhance the comparability between physical therapy practices. In 2020, an ad hoc task force of experts in the field of physical therapy, exercise therapy, rehabilitation science, respiratory medicine, general medicine and elderly care medicine, as well as patient representatives, developed a profiling system (the "Dutch model") for patients with COPD to allocate patients into subgroups for exercise-based care [36]. The participants in the current study suggested that baseline measures and patient characteristics needed for allocating patients to subgroups according to the profiling system should be included as process indicators in the core set. They stated that the stratification of patients into subgroups based on these profiles would enhance the comparability between practices.

Conditions for interpreting outcomes

Another reported problem was the limited amount of outcome data provided in the study, especially for the HHD. A possible reason could be that real-world data *via* national data registries were used for the data collection. These registries used predefined technical specifications [18]. During the conduct of the study, the HHD was a new measure implemented in the data registries and this may potentially have resulted in the low amount of provided data, which was also mentioned in the focus groups. Participants therefore suggested that the implementation of process measures is needed to stimulate routine data collection as a first step in quality improvement. When comparing outcomes, the participants were interested in the background information of patients with COPD, such as smoking status, exacerbations and body weight, for better interpretation of the differences in, for example, the change or end scores. When using these outcomes as a learning tool, the education of physical therapists is needed to gain knowledge about interpretation of the outcomes. Furthermore, to enhance the comparability between practices, participants suggested including only outcomes of patients that were treated for ⩾3 months.

Including the percentage of a predicted value

Absolute outcomes were used to calculate the end and change scores for the 6MWT and HHD. The participants suggested that outcomes should be presented as percentages of predicted values based on reference data from the healthy population [37, 38]. These normative values are based on previous research and can be calculated according to gender, age and body weight.

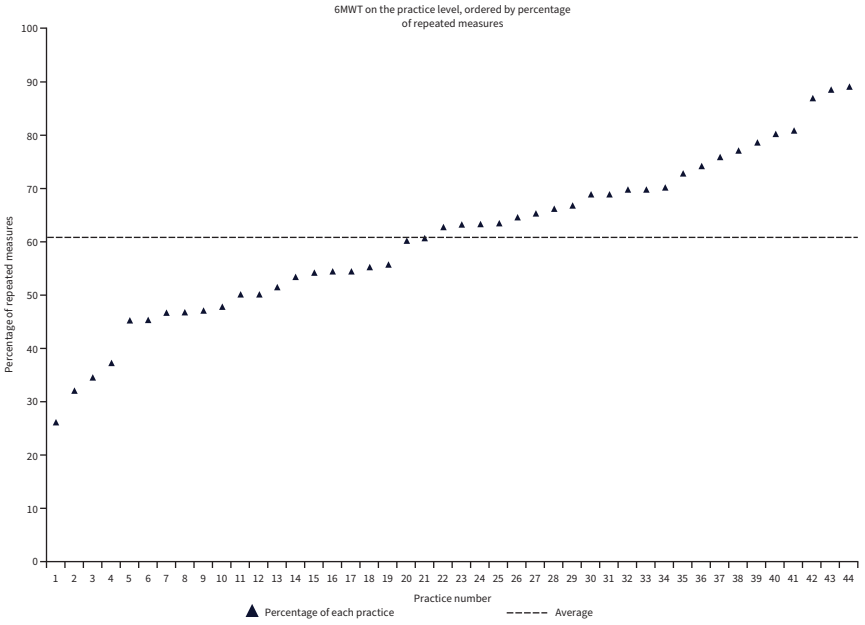
Selection of the core set

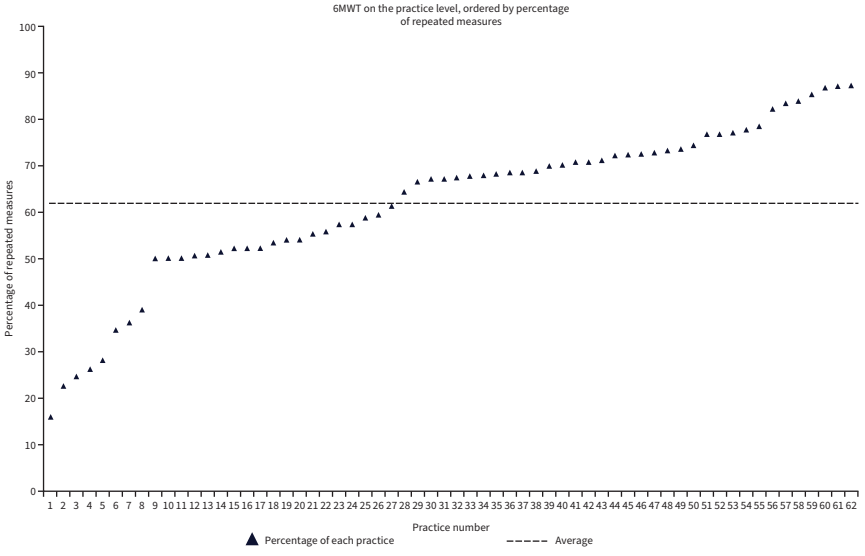
After discussing the outcomes, all (100%) of the physical therapists and senior researchers in each focus group preferred the inclusion of seven quality indicators in the core set: three process indicators for the routine measurement of the 6MWT, CCQ and HHD; three outcome indicators using the pre- to post-treatment change in the 6MWT, CCQ and HHD scores; and a combined process indicator to monitor

TABLE 5 Selected core set of quality indicators accepted by stakeholders based on the perceived added value as quality improvement tools

Type of indicator	Quality indicator description	Overall mean/ percentage[#]	Range[#]
Physical capacity measured with the 6MWT			
Process	The proportion of patients with COPD who underwent physical therapy treatment and who completed the 6MWT pre- and post-treatment to evaluate physical capacity	60.7%	26.1–88.8%
Outcome	The mean change score ±95% CI of patients with COPD who underwent physical therapy treatment and pre- and post-treatment measurement with the 6MWT to evaluate physical capacity	2.8 m	−5.4 to 13.4
Health-related quality of life measured with the CCQ			
Process	The proportion of patients with COPD who underwent physical therapy treatment and who completed the CCQ pre- and post-treatment to evaluate aspects of health-related quality of life	62.6%	14.8–88.7%
Outcome	The mean change score ±95% CI of patients with COPD who underwent physical therapy treatment and pre- and post-treatment measurement with the CCQ to evaluate health-related quality of life	–0.1	0.3 to –0.6
Quadriceps strength measured with the HDD			
Process	The proportion of patients with COPD who underwent physical therapy treatment and who completed the HHD pre- and post-treatment to evaluate quadriceps strength	31.4%	5.9–87.5%
Outcome	The mean change score ±95% CI of patients with COPD who underwent physical therapy treatment and pre- and post-treatment measurement with the HHD to evaluate quadriceps strength	7.5 Nm	2.7–13.1
Baseline measures for the 6MWT, CCQ, accelerometer and patient characteristics that can be used in a profiling system to stratify patients into subgroups for care[¶]			
Process	The proportion of patients with COPD who underwent physical therapy treatment and who completed the baseline measurements for the 6MWT, CCQ, accelerometer, gender, age, body weight and number of exacerbations in the past year	2.4%	

6MWT: 6-min walk test; CCQ: Clinical COPD Questionnaire; GPE-DV: Global Perceived Effect – Dutch Version; HHD: hand-held dynamometer; MRC: Medical Research Council Dyspnoea Scale; MCID: minimal clinically important difference. Only outcomes of patients that were included that were treated for ⩾3 months. [#]The overall mean/percentage and range are the outcomes of the physical therapy practices that provided ⩾10 cases, used for describing and selecting the quality indicators for the core set. [¶]Baseline measures and patient characteristics selected to allocate patients into subgroups based on the Dutch model for exercise-based care in primary care [36].


6MWT on the practice level, ordered by percentage of repeated measures
Percentage of repeated measures
Practice number
Percentage of each practice
Average


6MWT on the practice level, ordered by percentage of repeated measures
Percentage of repeated measures
Practice number
Percentage of each practice
Average

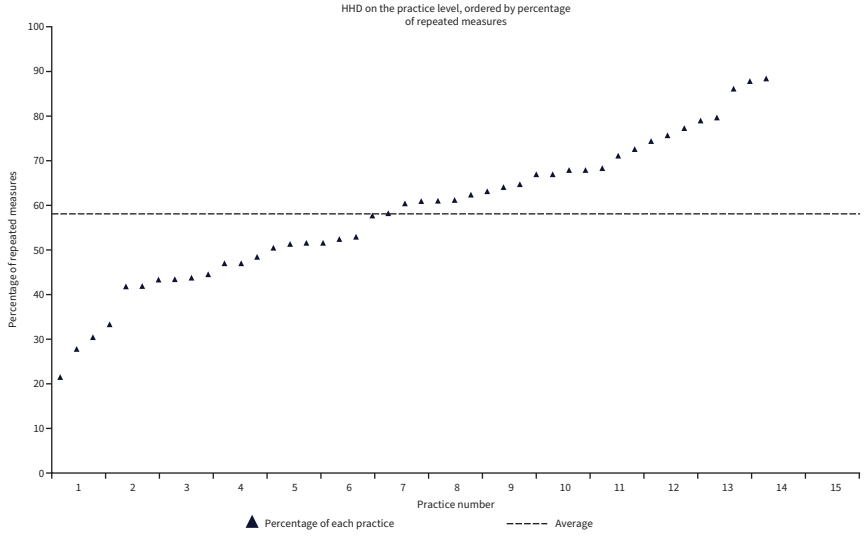

HHD on the practice level, ordered by percentage of repeated measures
Percentage of repeated measures
Practice number
Percentage of each practice
Average

FIGURE 1 a) The proportion of patients with COPD who underwent physical therapy treatment in which a pre- and/or post measure was provided. b) The proportion of patients with COPD who underwent physical therapy treatment in which a pre and/or post measure was provided for the CCQ. c) The proportion of patients with COPD who underwent physical therapy treatment in which a pre and/or post measure was provided for the HHD.

the baseline measurement of three measures (6MWT, CCQ and an accelerometer (steps per day)) and patient characteristics (age, gender, body weight and number of exacerbations in the past year) to allocate patients into subgroups based on the profiling system of the Dutch model [36]. The final core set of seven quality indicators is shown in table 5. Figure 1a (6MWT), 2b (CCQ) and 2c (HHD) present the proportion of patients with pre- and post-treatment Figures 2a (6MWT), 1b (CCQ) and 1c (HHD) presents the caterpillar plots of the quality indicators in the final core set.

Discussion

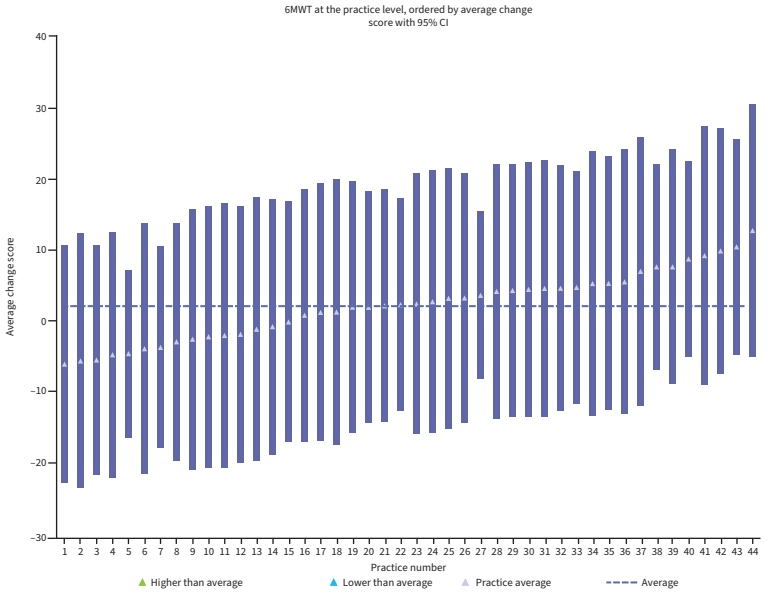
The major finding in this study is that all participants in the focus groups accepted the quality indicators as a quality improvement tool based on their perceived added value, and selected a core set of seven outcome-based quality indicators for patients with COPD. The final core set includes a process and outcome indicator for three outcomes: physical capacity measured with the 6MWT, health-related quality of life measured with the CCQ, and quadriceps strength measured with the HHD. A combined process indicator was included to monitor the baseline measurement of three measures used to allocate patients into subgroups based on the Dutch model profiling system [36]. To our knowledge, this is the first study to develop a core set of outcome-based quality indicators including a practice test for patients with COPD in physical therapy primary care practice. Using the core set, it is possible to compare standardised outcomes for patients between practices.

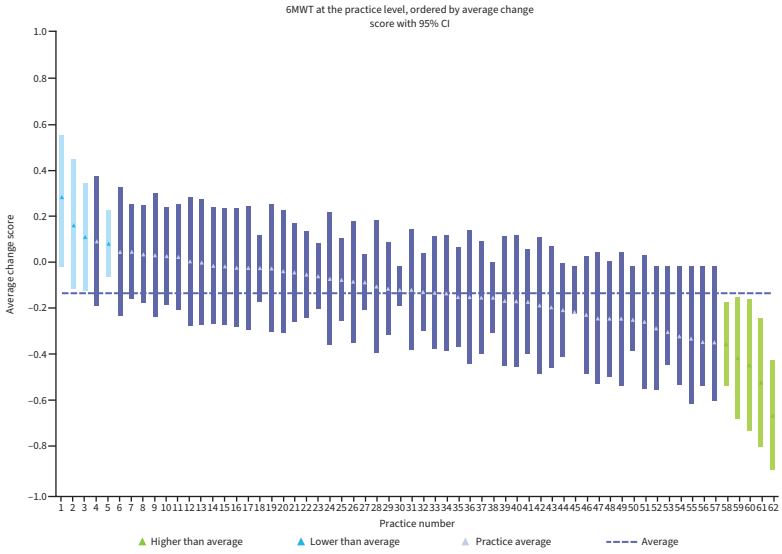
Several studies have developed quality indicators for COPD care [39–44], but most sets were developed to evaluate processes or structures of care, *e.g.* monitoring the proportion of patients for whom smoking status or the availability of exercise equipment was recorded [39, 41–44]. These studies differed in their care focus areas, which were hospitalised care, end-of-life care, transitional care after hospitalisation, pulmonary rehabilitation, vulnerable elders, or primary care in general [39–44]. None of these publications performed a practice test. In one indicator set, developed for pulmonary rehabilitation, some similar domains (physical capacity, strength and health-related quality of life) and measures (6MWT) were described [43]. The selection of change scores in the core set and the use of caterpillar plots is in line with other research describing the development of quality indicators based on PROMs [2, 5, 29]. A difference is that in the current study we focused specifically on development quality indicators based on outcomes of care, while other studies are more focused on evaluating processes of care [7, 45–49]. None of these studies aimed to develop a core set of outcome-based quality indicators to be used as quality improvement tools for healthcare providers, however. Quality indicators can also be developed for pay-for-performance initiatives, policy reports, insight into practice variation/delivered care, or the identification of differences in delivered care.

Despite the fact that our core set was developed in a Dutch environment, physical therapists in other countries could potentially use the indicator set in their daily practice. Nonetheless, the context of each country needs to be taken into account, specifically cultural or clinical practice differences between countries, such as differing guidelines or educational levels of physical therapists [50].

A strength of the current study is that we used a standard set of outcome domains and associated measures. The standard set was explicitly developed for patients with COPD being treated in primary care physical therapy practices, which was designed to be used as a basis for the further development of quality indicators [4]. The standard set is based on recommendations in guidelines and the supporting literature, and was selected in a RAND/UCLA Delphi procedure, which is one of the most common methods for the development of quality indicators [11, 51].

Another strength of our study is that we collected real-world data to perform a practice test prior to the selection of the core set, which was judged to be an essential step in evaluating the validity, reliability and feasibility of the indicators [9]. The interpretation of the practice test was discussed with end-users and guideline developers in focus groups. Including stakeholders in the development process is an important step for the successful implementation of quality indicators [51]. In the current study, we explicitly focused on the development of an indicator set for learning and quality improvement purposes for physical therapists. When quality indicators are designed for other purposes, such as a support tool for patients to


6MWT at the practice level, ordered by average change score with 95% CI
Average change score
Practice number
Higher than average
Lower than average
Practice average
Average


6MWT at the practice level, ordered by average change score with 95% CI
Average change score
Practice number
Higher than average
Lower than average
Practice average
Average

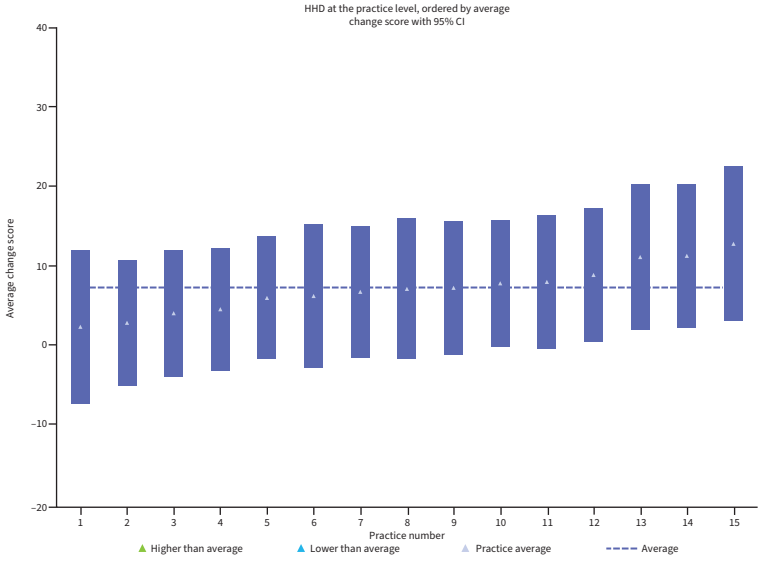

HHD at the practice level, ordered by average change score with 95% CI
Average change score
Practice number
Higher than average
Lower than average
Practice average
Average

FIGURE 2 a) The mean change score on the 6MWT with 95% CI of patients with COPD between pre- and post-physical therapy treatment. b) The mean change score on the CCQ with 95% CI of patients with COPD between pre- and post-physical therapy treatment. c) The mean change score on the HHD with 95% CI of patients with COPD between pre- and post-physical therapy treatment

choose providers, future research should also include other stakeholders (*i.e.* patients, policy-makers and health insurers) to evaluate their usefulness in daily practice.

A limitation of our study is that in multilevel analyses, a general rule of the thumb for the calculation of outcomes is the 30/30 rule (*i.e.* 30 physical therapy practices including a minimum of 30 patients each), allowing a valid comparison of indicator scores between practices [20, 21]. We did not use this rule of thumb as the threshold for estimating the case-mix adjusted scores for each quality indicator. The routine collection of clinical data by Dutch physical therapists treating patients with COPD is still in its infancy; therefore, we concluded that the 30/30 rule would not have been achievable in our study. Here, the collected data were only used as a supporting tool for the selection of the core set, so we decided to include physical therapy practices that had ⩾10 patients with COPD.

It is important to note that many practices did not reach the threshold of providing measurements for ⩾10 patients with COPD. When the process indicators, as presented in table 5, were based on all participating practices, the proportion of repeated measures was 39% for the 6MWT, 52% for the CCQ and 5% for the HHD. In our view, future implementation strategies must be conducted to improve the amount of data provided; for example, by giving feedback to practices with process indicators as presented in table 5.

Furthermore, due to the amount of data provided, we chose to compare the outcomes between physical therapy practices and not between physical therapists. When the amount of available data increases, there will be an opportunity to compare outcomes between physical therapists, both between and within practices. When sufficient data within practices are provided, physical therapists are able to learn from their own outcomes in comparison with peers who are employed in the same practice. We expect that, when comparing outcomes between physical therapists, the variability will be greater than between practices.

Another limitation is that we were not able to collect data that allowed us to allocate patients into subgroups based on their burden of disease, physical activity and physical capacity [4]. Hypothetically, the comparability and discriminability of the quality indicators would increase when allocating patients into subgroups. The participants of the focus groups underlined this hypothesis and suggested the inclusion of the Dutch profiling system for patients with COPD in the core set [36]; however, the Dutch profiling system had not yet been developed at the start of the data collection for this study. Future research could evaluate the core set for each subgroup to compare more homogeneous patient groups on their baseline characteristics. Another aspect to increase the comparability is to include more patient characteristics for case-mix adjustment. As suggested by patients with COPD and physical therapists, potentially relevant case-mix variables are, for example, smoking history, comorbidities and number of exacerbations [4].

Implications for practice

Outcome-based quality indicators based on real-world data, as provided in this study, can be used as a learning tool by comparing the collected patient outcomes between physical therapists or practices. This can, for example, be accomplished by discussing outcomes in peer assessment meetings of physical therapists to improve the quality of care. In such meetings, physical therapists critically appraise their peers' performance and give them constructive feedback [52–54]. In our opinion, Dutch physical therapists treating patients with COPD should first focus on expanding the amount of data collected. Giving feedback information can help to stimulate physical therapy practices to increase data collection. When sufficient data are provided and the comparison of outcomes in patient subgroups is established, the usability of the core set will increase. Future research should focus on the development of methods to improve the use of outcomes between peers and to set up specific actions to improve the quality of care.

Conclusion

This is the first study to describe and select a core set of seven outcome-based quality indicators for patients with COPD treated in primary care physical therapy practice. This core set includes process and outcome indicators related to measuring physical capacity, health-related quality of life and quadriceps strength, and a process measure for profiling patients within subgroups. To further evaluate the core

outcome set, future research should explore different strategies to promote data collection, including providing feedback of the outcomes to physical therapists.

Acknowledgements: We acknowledge our research members Juliette Cruijsberg and Janine Liefers; senior researcher Harald Miedema; senior physical therapists Cor Zagers and Leendert Tissink; and all the physical therapists who participated in this study, and provided data and/or participated in the focus group meetings.

Provenance: Submitted article, peer reviewed.

The datasets used and/or analysed during the current study are available from the corresponding author upon reasonable request.

All procedures were conducted according to the Declaration of Helsinki. The study protocol was approved by the Medical Ethical Committee of Radboud University Medical Center (registration number 2019-5455). Informed consent was obtained from all participants included in the current study.

Author contributions: Concept/research design: A.C. Verburg, S.A. van Dulmen, H. Kiers and P.J. van der Wees; manuscript writing: A.C. Verburg, S.A. van Dulmen, H. Kiers and M. Nijhuis-van der Sanden; data collection: A.C. Verburg, S.A. van Dulmen and H. Kiers; data analysis: A.C. Verburg; project management: A.C. Verburg and P.J. van der Wees; funding procurement: S.A. van Dulmen and H. Kiers; providing participants: A.C. Verburg, S.A. van Dulmen and H. Kiers; providing facilities/equipment: A.C. Verburg, S.A. van Dulmen and H. Kiers; providing institutional liaisons: A.C. Verburg, S.A. van Dulmen, H. Kiers and P.J. van der Wees; clerical/secretarial support: A.C. Verburg; consultation (including review of manuscript before submitting): S.A. van Dulmen, H. Kiers and P.J. van der Wees.

Conflict of interest: None declared.

Support statement: Funding support for this study was provided by Dutch health insurers CZ & the Friesland. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013; 346: f167.
2 Hildon Z, Neuburger J, Allwood D, *et al.* Clinicians' and patients' views of metrics of change derived from patient reported outcome measures (PROMs) for comparing providers' performance of surgery. *BMC Health Serv Res* 2012; 12: 171.
3 van der Wees PJ, Verkerk EW, Verbiest MEA, *et al.* Development of a framework with tools to support the selection and implementation of patient-reported outcome measures. *J Patient-Rep Outcomes* 2019; 3: 75.
4 Verburg AC, van Dulmen SA, Kiers H, *et al.* Development of a standard set of outcome domains and proposed measures for chronic obstructive pulmonary disease in primary care physical therapy practice in the Netherlands: a modified RAND/UCLA appropriateness method. *Int J Chron Obstruct Pulmon Dis* 2019; 14: 2649–2661.
5 Patient Reported Outcomes (PROs) in Performance Measurement Washington, DC: National Quality Form (NQF); 2012. Available from: http://www.qualityforum.org/Publications/2012/12/Patient-Reported_Outcomes_in_Performance_Measurement.aspx
6 Jette DU, Halbert J, Iverson C, *et al.* Use of standardized outcome measures in physical therapist practice: perceptions and applications. *Phys Ther* 2009; 89: 125–135.
7 Westby MD, Marshall DA, Jones CA. Development of quality indicators for hip and knee arthroplasty rehabilitation. *Osteoarthr Cartil* 2018; 26: 370–382.
8 Fitch K, Bernstein SJ, Aguilar MD, *et al.* The RAND/UCLA Appropriateness Method User's Manual. RAND Corporation. 2001.
9 Kötter T, Blozik E, Scherer M. Methods for the guideline-based development of quality indicators--a systematic review. *Implement Sci* 2012; 7: 21.
10 Vreeken HL, Beekman E, Doormaal van MCM, *et al.* Royal Dutch Society for Physical Therapy (KNGF) practice guideline COPD 2020.
11 Langendam MW, Piggott T, Nothacker M, *et al.* Approaches of integrating the development of guidelines and quality indicators: a systematic review. *BMC Health Serv Res* 2020; 20: 875.
12 Wollersheim H, Hermens R, Hulscher M, *et al.* Clinical indicators: development and applications. *Neth J Med* 2007; 65: 15–22.

13 Meerhoff GA, Verburg AC, Schapendonk RM, *et al.* Reliability, validity and discriminability of patient reported outcomes for non-specific low back pain in a nationwide physical therapy registry: a retrospective observational cohort study. *PLoS One* 2021; 16: e0251892.
14 Verburg AC, van Dulmen SA, Kiers H, *et al.* Patient-Reported outcome-based quality indicators in Dutch primary care physical therapy for patients with nonspecific low back pain: a cohort study. *Phys Ther* 2021; 101: pzab118.
15 Core Outcome Measures in Effectiveness Trials (COMET) Initiative 2019. Available from: https://www.comet-initiative.org/Studies/Details/1469
16 Gosselink R, Langer D, Burtin C, *et al.* KNGF-Guideline for physical therapy in patients with chronic obstructive pulmonary disease. https://www.fysionet-evidencebased.nl/images/pdfs/richtlijnen/copd_2008/copd_verantwoording_en_toelichting.pdf. 2008
17 Cuschieri S. The STROBE guidelines. *Saudi J Anaesth* 2019; 13 Suppl 1: S31–S34.
18 Meerhoff GA, van Dulmen SA, Maas MJM, *et al.* Development and evaluation of an implementation strategy for collecting data in a National Registry and the use of patient-reported outcome measures in physical therapist practices: quality improvement study. *Phys Ther* 2017; 97: 837–851.
19 Guideline protocol to standardize the testing procedure for patients with COPD (in Dutch). 2018.
20 Ali A, Ali S, Khan SA, *et al.* Sample size issues in multilevel logistic regression models. *PloS One* 2019; 14: e0225427.
21 Koolman X, Zuidgeest M, Visser J, *et al.* Indicatorstandaard 2.0: methodologische criteria voor de ontwikkeling van betrouwbare kwaliteitsindicatoren in de zorg.: Kwaliteitsinstituut; 2012.
22 Kessel PV, Triemstra M, de Boer D. Handreiking voor het meten van kwaliteit van zorg met Patient Reported Outcome Measures 2014. Available from: https://www.nivel.nl/sites/default/files/bestanden/Handreiking-meten-kwaliteit-zorg-met-Patient-Reported-Outcome-Measures.pdf
23 Greenhalgh J, Dalkin S, Gooding K, *et al.* Health Services and Delivery Research. Functionality and Feedback: a Realist Synthesis of the Collation, Interpretation and Utilisation of Patient-Reported Outcome Measures Data to Improve Patient Care. Southampton (UK), NIHR Journals Library, 2017.
24 Cazzola M, Hanania NA, MacNee W, *et al.* A review of the most common patient-reported outcomes in COPD–revisiting current knowledge and estimating future challenges. *Int J Chron Obstruct Pulmon Dis* 2015; 10: 725–738.
25 Twisk JWR. Applied Mixed Model Analysis: A Practical Guide. 2nd Ed. Cambridge, Cambridge University Press, 2019.
26 Gray J, Millett C, O'Sullivan C, *et al.* Association of age, sex and deprivation with quality indicators for diabetes: population-based cross sectional survey in primary care. *J R Soc Med* 2006; 99: 576–581.
27 Kiderlen M, Ponti A, Tomatis M, *et al.* Variations in compliance to quality indicators by age for 41 871 breast cancer patients across Europe: a European Society of Breast Cancer Specialists database analysis. *Eur J Cancer* 2015; 51: 1221–1230.
28 Nunes EV, Pavlicova M, Hu M-C, *et al.* Baseline matters: the importance of covariation for baseline severity in the analysis of clinical trials. *Am J Drug Alcohol Abuse* 2011; 37: 446–452.
29 Hildon Z, Allwood D, Black N. Making data more meaningful: patients' views of the format and content of quality indicators comparing health care providers. *Patient Educ Couns* 2012; 88: 298–304.
30 Hildon Z, Allwood D, Black N. Patients' and clinicians' views of comparing the performance of providers of surgery: a qualitative study. *Health Expect* 2015; 18: 366–378.
31 Boer D, Triemstra M, Koopman L, *et al.* Tool analyses; voor het vergelijken van patiëntervaringen tussen zorgaanbieders | versie 1.0 2016. Available from: https://www.zorginzicht.nl/binaries/content/documents/zorginzicht/ontwikkeltools/prom-toolbox/prom-cyclus-stap-5-definieren-indicator/prom-cyclus-stap-5-definieren-indicator/files/3/Tool-Analyse-voor-het-vergelijken-van-pati%C3%ABntervaringen-tussen-zorgaanbieders.pdf
32 Goldstein H, Michael JRH. The graphical presentation of a collection of means. *J R Stat Soc Ser A (Statistics in Society)* 1995; 158: 175–177.
33 Bohannon RW, Crouch R. Minimal clinically important difference for change in 6-minute walk test distance of adults with pathology: a systematic review. *J Eval Clin Pract* 2017; 23: 377–381.
34 Vaidya T, Beaumont M, de Bisschop C, *et al.* Determining the minimally important difference in quadriceps strength in individuals with COPD using a fixed dynamometer. *Int J Chron Obstruct Pulmon Dis* 2018; 13: 2685–2693.
35 Smid DE, Franssen FM, Houben-Wilke S, *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* 2017; 18: 53–58.
36 Spruit MA, Van't Hul A, Vreeken HL, *et al.* Profiling of patients with COPD for adequate referral to exercise-based care: the Dutch model. *Sports Med* 2020; 50: 1421–1429.
37 Troosters T, Gosselink R, Decramer M. Six minute walking distance in healthy elderly subjects. *Eur Respir J* 1999; 14: 270–274.

38 Andrews AW, Thomas MW, Bohannon RW. Normative values for isometric muscle force measurements obtained with hand-held dynamometers. *Phys Ther* 1996; 76: 248–259.

39 De Schreye R, Houttekier D, Deliens L, *et al.* Developing indicators of appropriate and inappropriate end-of-life care in people with Alzheimer's disease, cancer or chronic obstructive pulmonary disease for population-level administrative databases: a RAND/UCLA appropriateness study. *Palliat Med* 2017; 31: 932–945.

40 Michael F, Whitelaw S, Van Spall HG. Transitional care quality indicators to assess quality of care following hospitalisation for chronic obstructive pulmonary disease and heart failure: a systematic review protocol. *BMJ Open* 2019; 9: e032764.

41 Weidinger P, Nilsson JL, Lindblad U. Adherence to diagnostic guidelines and quality indicators in asthma and COPD in Swedish primary care. *Pharmacoepidemiol Drug Saf* 2009; 18: 393–400.

42 Gershon AS, Mecredy GC, Aaron SD, *et al.* Development of quality indicators for chronic obstructive pulmonary disease (COPD): a modified RAND appropriateness method. *Can J Respir Crit Care Sleep Med* 2019; 3: 30–38.

43 Dechman G, Cheung W, Ryerson CJ, *et al.* Quality indicators for pulmonary rehabilitation programs in Canada: a Canadian Thoracic Society expert working group report. *Can J Respir Crit Care Sleep Med* 2019; 3: 199–209.

44 Kleerup E. Quality indicators for the care of chronic obstructive pulmonary disease in vulnerable elders. *J Am Geriatr Soc* 2007; 55 Suppl 2: S270–S276.

45 Petersson IF, Strombeck B, Andersen L, *et al.* Development of healthcare quality indicators for rheumatoid arthritis in Europe: the eumusc.net project. *Ann Rheum Dis* 2014; 73: 906–908.

46 Peter WF, Hurkmans EJ, Wees PJvd, *et al.* Healthcare quality indicators for physiotherapy management in hip and knee osteoarthritis and rheumatoid arthritis: a Delphi study. *Musculoskeletal Care* 2016; 14: 219–232.

47 Scholte M, Neeleman-van der Steen CW, Hendriks EJ, *et al.* Evaluating quality indicators for physical therapy in primary care. *Int J Qual Health Care: J Int Soc Qual Health Care* 2014; 26: 261–270.

48 Oostendorp RA, Rutten GM, Dommerholt J, *et al.* Guideline-based development and practice test of quality indicators for physiotherapy care in patients with neck pain. *J Eval Clin Pract* 2013; 19: 1044–1053.

49 Oostendorp RAB, Elvers JWH, van Trijffel E, *et al.* Relationships between context, process, and outcome indicators to assess quality of physiotherapy care in patients with whiplash-associated disorders: applying Donabedian's model of care. *Patient Prefer Adherence* 2020; 14: 425–442.

50 Marshall MN, Shekelle PG, McGlynn EA, *et al.* Can health care quality indicators be transferred between countries? *Qual Saf Health Care* 2003; 12: 8–12.

51 Campbell SM, Braspenning J, Hutchinson A, *et al.* Research methods used in developing and applying quality indicators in primary care. *BMJ* 2003; 326: 816–819.

52 Maas MJ, Nijhuis-van der Sanden MW, Driehuis F, *et al.* Feasibility of peer assessment and clinical audit to self-regulate the quality of physiotherapy services: a mixed methods study. *BMJ Open* 2017; 7: e013726.

53 Maas MJM, Driehuis F, Meerhoff GA, *et al.* Impact of self- and peer assessment on the clinical performance of physiotherapists in primary care: a cohort study. *Physiother Can* 2018; 70: 393–401.

54 Maas MJ, van der Wees PJ, Braam C, *et al.* An innovative peer assessment approach to enhance guideline adherence in physical therapy: single-masked, cluster-randomized controlled trial. *Phys Ther* 2015; 95: 600–612.