Journal of Clinical Epidemiology 125 (2020) 130−137

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

# Primary care functioning scale showed validity and reliability in patients with chronic conditions: a psychometric study

Simone A.E. Postma[a,*], Henk Schers[a], Jules L. Ellis[b,c], Kees van Boven[a], Huib ten Napel[a,c], Hugo Stappers[a], Tim C. olde Hartman[a], Debby L. Gerritsen[a]

[a]*Radboud Institute for Health Sciences, Department of Primary and Community Care, Radboud University Medical Center, Nijmegen, The Netherlands*
[b]*Institute of Psychology, Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands*
[c]*WHO Collaborating Centre for Family of International Classifications, Bilthoven, The Netherlands*

Accepted 13 May 2020; Published online 29 May 2020

## Abstract

**Objectives:** We evaluated the psychometric properties of a newly developed self-report questionnaire that aims for a more person-centered approach in primary care for patients with chronic conditions, the Primary Care Functioning Scale (PCFS).

**Study Design and Setting:** To test the measurement properties of the PCFS, we asked patients with diabetes, cardiovascular disease, and chronic pulmonary disease to complete the PCFS questionnaire. The PCFS is entirely based on the International Classification of Functioning, Disability, and Health (ICF), consisting of 52 ICF-related items covering body functions, activities and participation, environmental factors, and personal factors. We analyzed three hypotheses representing different item sets of the 34 ICF-related items that assess the level of functioning (body functions, activities, and participation). We tested for unidimensionality, differential item functioning, reliability, and criterion-related validity.

**Results:** Five hundred and eighty-two patients completed the questionnaire. The total scores of the polytomous and dichotomized items from the overall set 'body functions, activities and participation' demonstrated unidimensionality, good reliability ($>0.80$), and stability over time without bias from background variables.

**Conclusion:** In sum, the PCFS can be used as a valid and reliable instrument to measure functioning in patients with chronic morbidity in primary care. 

*Keywords:* International classification of functioning; Disability and health; Primary care; Multimorbidity; Chronic morbidity; Psychometrics; Item response theory

## 1. Introduction

The management of patients with chronic conditions in the primary health care setting is a well-known challenge. For patients with chronic (multi)morbidity, general practitioners (GPs) balance between delivering care on the basis of a single-disease approach and providing more holistic and person-centered care [1−3]. Most primary care guidelines focus on the diagnosis and treatment of a specific single disease [4], and integrated guidelines for patients with more than one chronic condition and multimorbidity are lacking. Moreover, different disease-specific guidelines may contradict each other, which makes evidence-based care for patients with chronic conditions and multimorbidity even more complicated [5−9]. Patients with chronic conditions often seek contact with health care providers, and they report reduced health related daily functioning as a direct consequence of their multimorbidity [10−13]. So far, efforts to improve health outcomes for patients with multimorbidity are disappointing that may be due to a disease-oriented approach which generally does not take daily functioning into account [14−16]. A real person-centered approach would focus on what is important to the person, and this implies that the management of the changing symptoms is integrated within an individual's personal and social context.

The International Classification of Functioning, Disability and Health (ICF) has been developed to offer a

Declarations of interests: None.

Funding: This work was supported by ZonMw (grant number 839110017).

* Corresponding author. Department of Primary and Community Care, Postal Route 117PO Box 9101, 6500 HB Nijmegen, The Netherlands. Tel.: +31-24-36-16336; fax: +31-24-35-41862.

*E-mail address:* simone.postma@radboudumc.nl (S.A.E. Postma).

https://doi.org/10.1016/j.jclinepi.2020.05.018

**What is new?**

**Key findings**

- We assessed the psychometric properties of a primary care functioning scale concerning the measurement of the level of functioning on a sample of 565 patients.
- The total scores of the dichotomized items from body functions, activities and participation proved to be unidimensional, reliable and stable without bias of background variables.

**What this adds to what was known?**

- The primary care functioning scale is a valid and reliable instrument to measure the different aspects of functioning in patients with chronic morbidity in primary care.

**What are the implications and what should change now?**

- Including different aspects of functioning on the physical, mental and social domain in conversations in the clinical setting is essential for personalized healthcare.
- The primary care functioning scale can be used as a communication tool to cover all these aspects and supports the GP to objectify and monitor the level of functioning of patients.

foundation for defining and assessing health and health-related functioning within the individual's personal context. The ICF framework includes different components of functioning, such as body functions, activities and participation, environmental factors, and personal factors [17]. Until now, the ICF has not been used in the primary care setting and GPs are neither trained nor equipped with instruments to systematically assess the level of functioning. Recently, we developed a self-report questionnaire for patients with chronic conditions in primary care, the Primary Care Functioning Scale (PCFS), which assesses their level of functioning. It covers the aspects of the entire ICF framework, including personal factors [18]. In the development study, described elsewhere [18], we used the methodology of cognitive interviewing to match patients' responses to the ICF-based questions with the ICF itself. The PCFS has demonstrated high content validity and promising construct validity, that is, patients were well capable of describing their own health status in terms of ICF, and has the potential to support interprofessional communication in primary care, monitoring, registration, and classification at the level of a patient's functioning [18].

During consultations in primary care, the results of the PCFS can be used to enhance the conversation about the functional status of the patient and to help identify what is important to the patient. The ICF explicitly tries to move away from a unidirectional model that interprets functional status as a consequence of disease or multimorbidity. In primary care, the focus should also be on identifying opportunities for activities and participation that might improve body functions, and the PCFS might be able to identify these. In the clinical and research setting of primary care, it could also be highly relevant to measure the level of functioning of patients with chronic conditions and multimorbidity, and moreover give opportunities to evaluate the effectiveness of interventions in terms of functioning, next to disease-specific clinical outcomes. Therefore, the aim of this study is to assess the psychometric properties (i.e., validity and reliability) of the PCFS concerning the measurement of the level of functioning on a sample of patients with chronic conditions in primary care.

## 2. Methods

### 2.1. Participants and design

This psychometric study was carried out in 15 primary care practices located in different urban and rural areas within the practice-based research network Nijmegen (PBRN), from January 2017 to July 2017. Patients who visited the practice for a routine check-up with the practice nurse for diabetes mellitus (DM), chronic obstructive pulmonary disease (COPD)/asthma, or cardiovascular risk management were invited by their practice nurse to participate. Twenty-six practice nurses from 15 primary care practices were asked to recruit 25 patients each. The inclusion criteria for participation were as follows: age 50 years or older, the presence of DM 1 or 2, COPD/asthma or a cardiovascular disease, and the ability to complete a written questionnaire. Patients were asked to complete the PCFS questionnaire at home and return the completed questionnaire to the researchers in a stamped envelope. Patients could also indicate if they were willing to complete the questionnaire for a second time (retest) approximately 2 weeks after completion of the questionnaire. Patients did not receive any incentive to participate. In case of nonresponse, the practice nurse collected demographic information (age, sex, and indication for routine check-up with practice nurse) and reasons for nonresponse. The study was carried out in accordance with Dutch privacy legislation. The Research Ethics Committee Arnhem/Nijmegen gave approval to conduct the study (registration number 2016-2499).

Written informed consent was obtained from all participating patients; patients were able to withdraw their consent at any time.

### 2.2. Measurement instrument

The PCFS instrument is a self-report questionnaire. It consists of 52 ICF-related items corresponding to ICF's

components *body functions* (items 1−11), *activities and participation* (items 12−34), *environmental factors* (items 35−44), and *personal factors* (items 45−52) (see Appendix 1) [18]. The 34 items from the components body functions and activities and participation assess the actual level of functioning, whereas the 18 items from the components environmental factors and personal factors are linked to the level of functioning. We will analyze the psychometric properties of the 34 items from body functions and activities and participation that assess the actual level of functioning. Patients are asked to rate how they feel at this moment using response scales ranging from *no problem* (score 1) to *complete problem* (score 5), with a higher score representing decreased functioning.

### 2.3. Statistical analysis

We formed three related hypotheses about how the 34 items' responses from ICF component's body functions and activities and participation could be summarized in a total score per participant (unidimensionality), without bias of background variables (no differential item functioning) that reliably discriminates between the participants. Hypothesis 1 poses that all items form one dimension, the pertaining item set is called set 1 (body functions, activities and participation: items 1−34). Because it was not clear in advance whether all items could be summarized in a total score, we formed a second and third hypothesis. Hypothesis 2 supposes two existing dimensions with corresponding item sets named set 2 (body functions: items 1−11) and set 3 (activities and participation: items 12−34). Hypothesis 3 states that there are three dimensions, including set 2 (body functions: items 1−11), set 4 (activities: items 12−27), and set 5 (participation: items 28−34). This is illustrated in Figure 1.

#### 2.3.1. Unidimensionality

To assess the unidimensionality of the PCFS instrument (i.e., whether the items of the questionnaire measure the same single construct of interest), we used the monotone polynomial model [19]. We tested the item sets for polytomous response categories (1 = no problem, 2 = mild problem, 3 = moderate problem, and 4 = severe and complete problem) and dichotomized response categories (1 = no problem and 2 = mild to complete problem). To increase statistical power, the highest response categories (severe and complete problem) were collapsed into a single category. For the same reason, we also tested the models for dichotomized items. Items 21 and 28−34 (see Appendix 1) have a responding category 'not applicable', and it was not immediately clear whether this category should be recoded to the low value (no problem) or to the high value (mild to complete problem). Recoding this category into both options for dichotomized items revealed that the fit was best when recoding to 'no problem' (data not shown). Therefore, we recoded the category 'not applicable' into 'no problem' throughout all further analyses. Items 31, 32, and 33 (see Appendix 1) are more or less mutually exclusive; therefore, these items were combined into a single variable holding the maximum of these items. This variable will be named 'item 35' in the results. Participants with missing values for items 1−34 were removed from the analyses.

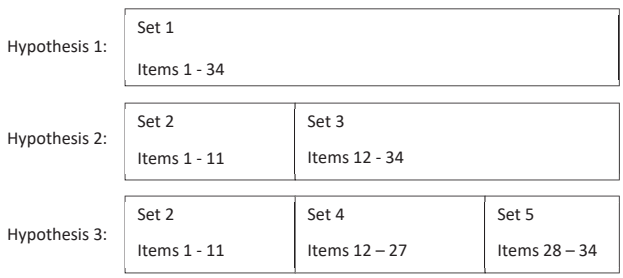


**Fig. 1.** Hypotheses of the investigated item sets.

The fit with the model was assessed with the R package for multidimensional item response theory (MIRT) [20].

Each of the item sets was tested for unidimensionality with a monotone polynomial model with polynomials of degree 3. We based our decisions on the *P*-value and the root mean square error approximation (RMSEA) of the $M_2$-test [21]. $M_2$ is an overall model fit test and the RMSEA derived from it can be considered as an effect size estimate. We interpret the fit of the model as *good* if $P > 0.05$ or RMSEA $\leq 0.05$ and as *acceptable* if $P \leq 0.05$ and RMSEA $> 0.05$ and the lower bound of the 90% confidence interval of the RMSEA is less than or equal to 0.05. In all other cases, the fit of the model is considered *unacceptable*.

#### 2.3.2. Differential item functioning

Differential item functioning (DIF) analysis was conducted to explore whether items of the questionnaire operate in the same way in different subgroups of individuals (e.g., age, sex, and education). Individuals from different subgroups with the same level of functioning should have the same probability of giving a certain expected response on an item, and the existence of DIF in an item could indicate that an item is biased [22]. DIF was explored for the unidimensional sets, and the selected external variables were dichotomized with a cutoff score. The external variables, with their cutoff scores between parentheses, were as follows: sex, age (score 1: $\leq$67 years, score 2: $>$67 years), education (score 1: $\leq$ intermediate general secondary education, score 2: $>$ intermediate general secondary education), number of prescribed drugs (score 1: $\leq$3 prescribed drugs, score 2: $>$3 prescribed drugs), and number of reported chronic conditions (score 1: $\leq$2 chronic conditions, score 2: $>$2 chronic conditions). We tested DIF in the way outlined by Meade [23]. This means that the analysis starts with the fully constrained model and that items are dropped if they show DIF. Next, for the variables that demonstrated significant DIF with $P < 0.05$, the effect sizes were estimated, whereas the other items that

demonstrated no DIF were used as anchor. As a measure of effect size, we used Dmax, which expresses the effect of DIF on the scale of the test score. Because the items are dichotomized, the maximum score equals the number of items, which is expressed as J. The ratio will be called the relative effect size (RES) = |Dmax|/J, where J is the number of items in the set. We will regard the DIF as substantial if RES >0.05.

#### 2.3.3. Reliability

We assessed the reliability of the unidimensional sets with two different methods. In the first method, we derived the reliability from the estimated model parameters with the 'marginal_rxx' method of MIRT. In the second method, we assessed the reliability with the traditional Cronbach's alpha. Test–retest reliability was determined for the mean scores of the unidimensional sets by calculating the intraclass correlation for absolute agreement of single measures. This coefficient can be interpreted as a measure of stability of the scores. We excluded 35 patients from the test–retest reliability analysis because of reported changes in their functional health status between the first and second test. Reliability coefficients will be labeled as sufficient ($0.70 < r < 0.80$) or good ($r > 0.80$) or insufficient ($r < 0.70$) [24]. All reliability analyses were performed on dichotomized items. Only participants who answered at least 80% of the items per set were included.

#### 2.3.4. Criterion-related validity

Apart from the PCFS, patients rated general questions about their experienced health that were related to the scores of the PCFS. For example, patients were asked how they would rate their physical health, mental health, and quality of life during the past month on five point Likert scale from very good to poor (see Appendix 1). In this case, the test scores served as dependent variables in an ANOVA with the ratings as an independent categorical variable (fixed factor).

## 3. Results

### 3.1. Patient characteristics

We asked 708 patients from 15 general practices to participate in the study. In total, 582 patients (82%) consented and completed the PCFS questionnaire. We excluded 17 patients from the sample because of exclusion criteria (age < 50 years). Nonresponse analysis revealed that patients did not want to participate because of experienced time constraints, participation in other (previous) studies, no interest, or other individual reasons. Among the nonresponding patients were significantly more women (53.7%) than the responding patients (41.1%). Sociodemographic and clinical characteristics of the study participants ($n = 565$) were determined before further analyses and are

**Table 1.** Sociodemographic characteristics of the study participants ($n = 565$)

| Variables | *n* (%) |
|---|---|
| Sex | |
| Male | 333 (58.9) |
| Female | 232 (41.1) |
| Mean age, years (SD) | 68.5 |
| Current marital status | |
| Married or in a relationship | 427 (75.6) |
| Single, divorced, or Widowed | 137 (24.2) |
| Level of education[a] | |
| Low | 204 (36.1) |
| Middle | 190 (33.6) |
| High | 137 (24.2) |
| Other | 30 (5.3) |
| Occupation | |
| Working (remunerative, nonremunerative) | 149 (26.4) |
| Nonworking (retired, incapacity to work) | 414 (73.3) |
| Morbidity | |
| Reported 0 (chronic) condition | 36 (6.4) |
| Reported at least one (chronic) condition | 529 (93.6) |
| Reported at least two (chronic) conditions | 382 (67.6) |
| Reported at least three (chronic) conditions | 206 (36.5) |
| Reported four or more (chronic) conditions | 116 (20.5) |
| Number of prescribed drugs | |
| 0–1 | 39 (6.9) |
| 2–3 | 171 (30.3) |
| >3 | 350 (61.9) |
| Self-rated physical health during the past month | |
| Very good to good | 273 (48.3) |
| Fair | 195 (34.5) |
| Mediocre to poor | 92 (16.3) |
| Self-rated mental health during the past month | |
| Very good to good | 419 (74.2) |
| Fair | 108 (19.1) |
| Mediocre to poor | 34 (6) |
| Self-rated overall quality of life | |
| Very good to good | 403 (71.3) |
| Fair | 131 (23.1) |
| Mediocre to poor | 26 (4.6) |

[a] Level of education was categorized into low (no education and primary and lower secondary education), middle (upper secondary education), and high (preuniversity, higher vocational training, and university). Missing values are not included in the percentage calculation.

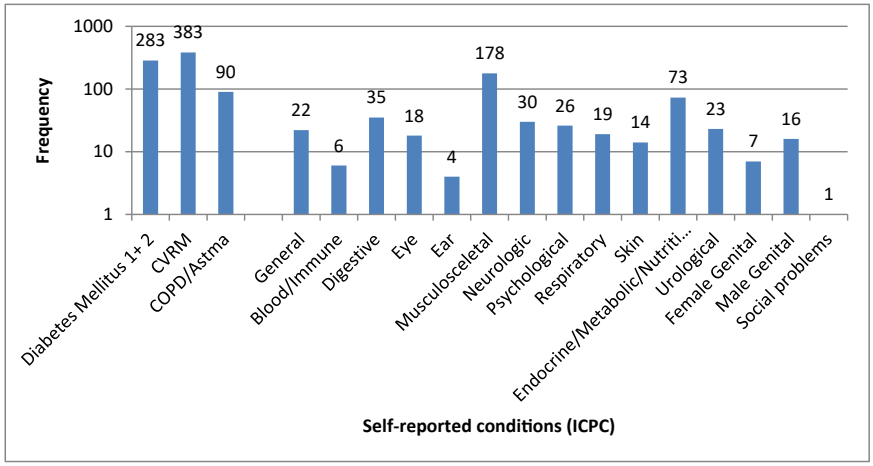


**Fig. 2.** Self-reported medical conditions of the study participants ($n$ = 565) in accordance with the chapters of the International Classification of Primary Care (ICPC).

presented in Table 1 and Fig. 2. At least 67.6% of the study participants reported to have two or more chronic morbidities. In Table 2, their overall level of satisfaction with functioning is reported.

### 3.2. Psychometric properties of PCFS

We decided to remove items 22–26 (see Appendix 1) from further analysis because of an extremely skewed response (>90% of response in extreme response category). Further statistical scale analyses were conducted for the three newly formed sets.

#### 3.2.1. Unidimensionality

After removing the participants with missing values, a total of 515 participants were used in the unidimensionality analyses. The results of the tests are presented in Table 3.

Four of the five sets displayed good fit for unidimensionality after dichotomizing the answers. This is less clear with polytomous items because this model could not be tested for set 2 (body functions), set 4 (activities), and set 5 (participation) (as the degrees of freedom were too small) and the RMSEA is 0.053 for set 1 (body functions, activities and participation) and 0.050 for set 3 (activities and participation). Therefore, we proceeded with the dichotomized items from set 1–3 displaying good unidimensionality (representing hypotheses 1 and 2) in our further analysis.

### 3.3. Differential item functioning

The items with DIF and relative effect size (RES) for set 1, 2, and 3 are presented in Table 4. Eight items of set 1 demonstrated DIF of age and/or sex, that is, these items did not operate similarly across different groups of sexes and ages, but the effect is small in accordance with the RES statistic. The RES statistic of 0.02 in age and sex means that age and sex caused a difference of 2% of the scale's range on average (this is illustrated in Appendix 2). There was substantial DIF of age in set 2; the significant differences were found in items 3, 6, 8, and 9. This means that for these items, there was an effect of age that was stronger or weaker than for the other items. There was also significant DIF of sex in set 2 and sex and age in set 3, but these effects were not substantial as assessed with RES.

**Table 2.** Satisfaction with overall level of functioning, environmental factors, and personal factors of the study participants ($n$ = 565)

| Satisfaction level | *n* (%) |
|---|---|
| Satisfaction overall level of body functions, mean | |
| Yes | 309.3 (54.7) |
| Neutral | 129.9 (22.9) |
| No | 121.2 (21.4) |
| Satisfaction overall level of activities, mean | |
| Yes | 404.7 (71.6) |
| Neutral | 87.3 (15.5) |
| No | 70.3 (12.4) |
| Satisfaction overall level of participation, mean | |
| Yes | 434.3 (76.9) |
| Neutral | 78.1 (13.8) |
| No | 36 (6.4) |
| Satisfaction overall environmental factors, mean | |
| Yes | 434.3 (76.9) |
| Neutral | 104.4 (18.5) |
| No | 15.3 (2.7) |
| Satisfaction overall personal factors, mean | |
| Yes | 418.3 (74) |
| Neutral | 129.4 (22.9) |
| No | 13 (2) |

Missing values are not included in the percentage calculation.

**Table 3.** Results of goodness-of-fit tests of unidimensional monotonic polynomial model of the five different item sets

| Set | Answer levels | Goodness-of-fit test | Fit |
|---|---|---|---|
| 1 | 2 | $M_2(270) = 589.22$, $P < 0.001$, RMSEA = 0.048, CI90 = [0.043, 0.053] | Good |
| 1 | 4 | $M_2(216) = 522.83$, $P < 0.001$, RMSEA = 0.053, CI90 = [0.047, 0.058] | Acceptable |
| 2 | 2 | $M_2(22) = 33.94$, $P = 0.05$, RMSEA = 0.032, CI90 = [0.001, 0.053] | Good |
| 2 | 4 | Not available (degrees of freedom too small) | Unknown |
| 3 | 2 | $M_2(72) = 152.4$, $P < 0.001$, RMSEA = 0.047, CI90 = [0.036, 0.057] | Good |
| 3 | 4 | $M_2(40) = 91.93$, $P < 0.001$, RMSEA = 0.050, CI90 = [0.037, 0.064] | Acceptable |
| 4 | 2 | $M_2(22) = 40.97$, $P = 0.008$, RMSEA = 0.041, CI90 = [0.02, 0.06] | Good |
| 4 | 4 | Not available (degrees of freedom too small) | Unknown |
| 5 | 2 | Not available (degrees of freedom too small) | Unknown |
| 5 | 4 | Not available (degrees of freedom too small) | Unknown |

### 3.4. Reliability

The reliabilities of item sets 1, 2, and 3 were good (>0.80, Table 5). The Cronbach's alpha coefficients were good for set 1 and 3 (>0.80) and sufficient for set 2 ($0.70 < r < 0.80$). The interval between the test and retest varied between two and a maximum of 7 weeks (excluding one outlier of 10 weeks). Despite the high test–retest reliabilities, the mean scores decreased significantly for set 1, 2, and 3 with 3%, 3%, and 2%, respectively.

#### 3.4.1. Criterion-related validity evidence

Number of diseases and ratings of physical health, mental health, and quality of life explained about 50% of the variance of the test scores of the PCFS (Table 6). Age was not related to the scores (R squared = 0.003, F(1, 513) = 1.556, $P = 0.213$).

**Table 4.** Items with DIF and relative effect sizes (RES) of DIF tests for set 1–3

| Set 1 | Items with DIF | RES |
|---|---|---|
| Sex | 6, 27, 35 | 0.02 |
| Age | 3, 6, 8, 14, 15, 19, 27, 35 | 0.02 |
| Education | - | |
| Medication | - | |
| Conditions | - | |
| Set 2 | | |
| Sex | 6 | 0.02 |
| Age | 3, 6, 8, 9 | **0.08** |
| Education | - | |
| Medication | - | |
| Conditions | - | |
| Set 3 | | |
| Sex | 27, 35 | 0.03 |
| Age | 14, 19, 35 | 0.03 |
| Education | - | |
| Medication | - | |
| Conditions | - | |

The effect sizes are in boldface if RES > 0.05

## 4. Discussion

### 4.1. Summary

We studied the measurement properties of the items body functions and activities and participation from the PCFS, a self-report questionnaire for patients with chronic conditions in the general practice population. The total scores of the polytomous and dichotomized items of the PCFS questionnaire corresponding with ICF's components body functions and activities and participation (set 1) proved to be unidimensional, reliable, and fairly stable over a short period of time without bias of background variables. The total scores of the PCFS corresponding with ICF's component body functions (set 2) and ICF's component activities and participation (set 3) separately were unidimensional for dichotomized items only, and this score seemed reliable and stable.

### 4.2. Strengths and limitations

A strength of this study is that it is unique in assessing the psychometric properties of an instrument measuring all different aspects of functioning tailored to patients with chronic morbidity in primary care. This is the first time that a self-report questionnaire directly developed from the ICF primary care set that is entirely based on the ICF has been validated [18]. Furthermore, an important strength of the present study is the high response rate of 82% and the large data set of 565 patients with chronic conditions in primary care

**Table 5.** Reliabilities in accordance with IRT model, Cronbach's alpha, and test–retest intraclass correlations for absolute agreement

| Item set | Model | Alpha | Test–retest |
|---|---|---|---|
| 1 | 0.95 | 0.89 | 0.91 ($n = 217$) |
| 2 | 0.87 | 0.76 | 0.83 ($n = 217$) |
| 3 | 0.91 | 0.84 | 0.89 ($n = 218$) |

**Table 6.** Variance of test score set 1–3 explained by 1) number of diseases, physical health, mental health, and quality of life ratings as between-participant factors separately and 2) number of diseases, physical health, mental health, and quality of life ratings as between-participant factors jointly

| Variables per item set | R squared in the test score | Adjusted R squared |
|---|---|---|
| Set 1 | | |
| Number of diseases | 0.178 | 0.171 |
| Physical health rating | 0.379 | 0.374 |
| Mental health rating | 0.246 | 0.240 |
| Quality of life rating | 0.291 | 0.286 |
| Jointly | 0.628 | 0.501 |
| Set 2 | | |
| Number of diseases | 0.140 | 0.132 |
| Physical health rating | 0.293 | 0.288 |
| Mental health rating | 0.221 | 0.216 |
| Quality of life rating | 0.215 | 0.209 |
| Jointly | 0.553 | 0.400 |
| Set 3 | | |
| Number of diseases | 0.170 | 0.162 |
| Physical health rating | 0.364 | 0.359 |
| Mental health rating | 0.212 | 0.206 |
| Quality of life rating | 0.286 | 0.281 |
| Jointly | 0.604 | 0.468 |

having completed the PCFS questionnaire, both increasing the precision, power, and reliability of the measurement properties of the instrument. This study also has some limitations that need to be addressed. First, our findings from the Dutch general practice context of patients with chronic conditions cannot be generalized to other samples or settings without further external validation studies. Second, we did not test the convergent validity, that is, if domains of the PCFS are related with other validated measurement instruments, for example the Short-form 36 Questionnaire (SF-36) or the World Health Organization Disability Assessment Schedule (WHODAS 2.0). The PCFS questionnaire consists of 52 items, which might be a burden for some patients to complete; therefore, we decided not to use additional questionnaires next to the PCFS in our study. However, we did use information from clinical questions and self-report ratings of physical health, mental health, and overall quality of life to assess criterion-related validity. Furthermore, the time between the test and retest interval varied. Preferably, the second test was conducted at a 2-week interval for every participant; however, in practice, this was not possible for patient-related and organizational reasons.

### *4.3. Comparison with the literature*

Previously, Cieza et al. found that it is possible to construct a measurement of functioning by using ratings of ICF categories by clinicians from an ICF set for patients with osteoarthritis [25]. In line with this finding, our study shows that the self-reported ratings of ICF-based items from patients with chronic conditions are suitable to measure the different aspects of functioning. Clinical measurements of functioning will provide the GP with a score of the overall level of functioning, which could be important in the complex management of the growing population of patients with multimorbidity in primary care. To our knowledge, the ICF-based items from the PCFS represent the only self-report measure of functioning for patients with chronic conditions in primary care that incorporates all the different aspects of functioning, including body functions and activities and participation. Until now, most measurement instruments only include a limited number of aspects of functioning. For example, the SF-36 includes scales for physical functioning and social functioning; however, these items do not cover the entire construct of functioning when compared with the different bio, psycho, and social perspectives that the World Health Organization uses for the conceptual basis when defining functioning [17].

### *4.4. Implications for research and/or clinical practice*

With the establishment of the PCFS as a measurement instrument that includes a comprehensive set of the different aspects of functioning, the PCFS can be used to objectify and monitor the functioning of patients with chronic conditions in primary care. For GPs, it is obvious that including aspects of functioning and context in conversations in the clinical setting is essential for personalized health care; the PCFS can be used as a communication tool to cover all these aspects. The items of body functions and activities and participation from the PCFS can be used as a clinical measurement tool, measuring the level of functioning that enables the GP to identify patients in a high variety of functional profiles ranging from normal functioning to severely decreased functioning on physical, mental, and social domains. This is crucial because the number of patients that suffers from chronic conditions in primary care is still growing. In sum, this study demonstrates the PCFS questionnaire is a valid and reliable instrument for primary care to measure the different aspects of functioning for dichotomized response categories, in patients with chronic conditions. Further research with the PCFS is needed to study whether the PCFS is also a feasible, efficient, and practical instrument for GPs and practice nurses.

## CRediT authorship contribution statement

**Simone A.E. Postma:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing - original draft, Writing - review & editing. **Henk Schers:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing - original draft, Writing - review & editing. **Jules L. Ellis:** Conceptualization, Data curation, Formal analysis,

Methodology, Resources, Software, Validation, Visualization, Writing - original draft, Writing - review & editing. **Kees van Boven:** Conceptualization, Investigation, Methodology, Project administration, Supervision, Writing - original draft, Writing - review & editing. **Hugo Stappers:** Conceptualization, Investigation, Methodology, Supervision, Writing - original draft, Writing - review & editing. **Tim C. olde Hartman:** Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing - original draft, Writing - review & editing. **Debby L. Gerritsen:** Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing.

## Acknowledgments

The authors thank all the patients for their cooperation in this study.

Authors' contributions for the manuscript 'The Primary Care Functioning Scale: A valid and reliable instrument for patients with chronic conditions' by Simone Postma, Henk Schers, Jules Ellis, Kees van Boven, Huib ten Napel, Hugo Stappers, Tim olde Hartman, and Debby Gerritsen.

## Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2020.05.018.

## References

[1] Guthrie B, Payne K, Alderson P, McMurdo ME, Mercer SW. Adapting clinical guidelines to take account of multimorbidity. Br Med J 2012;345:e6341.

[2] Grol R. Development of guidelines for general practice care. Br J Gen Pract 1993;43(369):146−51.

[3] Grimshaw J, Thomas R, MacLennan G, Fraser C, Ramsay C, Vale L, et al. Effectiveness 8and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004;8:iii−iv. 1−72.

[4] Tinetti ME, Fried T. The end of the disease era. Am J Med 2004; 116(3):179−85.

[5] Tinetti ME, Bogardus ST Jr, Agostini JV. Potential pitfalls of disease-specific guidelines for patients with multiple conditions. N Engl J Med 2004;351:2870−4.

[6] Fried TR, Tinetti ME, Iannone L. Primary care clinicians' experiences with treatment decision making for older persons with multiple conditions. Arch Intern Med 2011;171:75−80.

[7] Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. JAMA 2005;294:716−24.

[8] Lugtenberg M, Burgers JS, Clancy C, Westert GP, Schneider EC. Current guidelines have limited applicability to patients with comorbid conditions: a systematic analysis of evidence-based guidelines. PLoS One 2011;6:e25987.

[9] Van Weel C, Schellevis FG. Comorbidity and guidelines: conflicting interests. Lancet 2006;367(9510):550−1.

[10] Fortin M, Bravo G, Hudon C, Lapointe L, Almirall J, Dubois M-F, et al. Relationship between multimorbidity and health-related quality of life of patients in primary care. Qual Life Res 2006;15:83−91.

[11] Kadam U, Croft P. Clinical multimorbidity and physical function in older adults: a record and health status linkage study in general practice. Fam Pract 2007;24:412−9.

[12] Salisbury C, Johnson L, Purdy S, Valderas JM, Montgomery AA. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. Br J Gen Pract 2011;61(582):e12−21.

[13] Lee TA, Shields AE, Vogeli C, Gibson TB, Woong-Sohn M, Marder WD, et al. Mortality rate in veterans with multiple chronic conditions. J Gen Intern Med 2007;22:403.

[14] Hughes LD, McMurdo ME, Guthrie B. Guidelines for people not for diseases: the challenges of applying UK clinical guidelines to people with multimorbidity. Age Ageing 2013;42:62−9.

[15] Kenning C, Fisher L, Bee P, Bower P, Coventry P. Primary care practitioner and patient understanding of the concepts of multimorbidity and self-management: a qualitative study. SAGE Open Med 2013;1. 2050312113510001.

[16] Stange KC, Ferrer RL. The paradox of primary care. Ann Fam Med 2009;7(4):293−9.

[17] WHO. International classification of functioning, disability and health (ICF). Geneva, Switzerland: WHO; 2001: 2013.

[18] Postma SAE, van Boven K, Ten Napel H, Gerritsen DL, Assendelft WJJ, Schers H, et al. The development of an ICF-based questionnaire for patients with chronic conditions in primary care. J Clin Epidemiol 2018;103:92−100.

[19] Falk CF, Cai L. Maximum marginal likelihood estimation of a monotonic polynomial generalized partial credit model with applications to multiple group analysis. Psychometrika 2016;81(2):434−60.

[20] Chalmers RP. mirt: a multidimensional item response theory package for the R environment. J Stat Softw 2012;48(6):1−29.

[21] Maydeu-Olivares A, Joe H. Assessing approximate fit in categorical data analysis. Multivariate Behav Res 2014;49(4):305−28.

[22] Zumbo BD. Three generations of DIF analyses: considering where it has been, where it is now, and where it is going. Lang Assess Q 2007; 4(2):223−33.

[23] Meade AW. Guide to IRT invariance using the MIRT package in R 2016. https://meade.wordpress.ncsu.edu/files/2019/03/Guide_to_IRT_Invariance_Tests_in_R.pdf. Accessed March 22, 2019.

[24] Evers A, Sijtsma K, Lucassen W, Meijer RR. The Dutch review process for evaluating the quality of psychological tests: history, procedure, and results. Int J Test 2010;10(4):295−317.

[25] Cieza A, Hilfiker R, Chatterji S, Kostanjsek N, Ustun BT, Stucki G. The International Classification of Functioning, Disability, and Health could be used to measure functioning. J Clin Epidemiol 2009;62:899−911.