RHEUMATOLOGY

Rheumatology 2013;52:924–932
doi:10.1093/rheumatology/kes401
Advance Access publication 18 January 2013

Original article

Psychometric properties of the Norwegian version of the Patient Generated Index in patients with rheumatic diseases participating in rehabilitation or self-management programmes

Mari Klokkerud[1,2], Margreth Grotle[1,3], Ida Løchting[1,3], Ingvild Kjeken[1,2], Kåre Birger Hagen[1,2] and Andrew M. Garratt[1,4]

Abstract

Objective. In rehabilitation, treatment is individually tailored to each patient's goals. Individualized instruments allow patients to choose domains that they consider important, which may make them particularly appropriate as evaluative tools in this setting. We aimed to evaluate the psychometric properties of the Norwegian version of the Patient Generated Index (PGI) in patients with rheumatic diseases participating in inpatient rehabilitation or self-management programmes.

Methods. Patients completed the PGI together with other outcome measures at arrival and 5 and 52 weeks after arrival. The PGI was assessed for data quality by completion rates, reliability by the intraclass correlation coefficient (ICC), agreement by standard error of measurement (SEM) and smallest detectable change (SDC). Construct validity was assessed by testing *a priori* hypotheses regarding correlation between PGI scores and other outcome measures. Responsiveness was assessed by an *a priori* hypothesis regarding the correlation of different change scores and standardized response means (SRMs).

Results. A total of 145 patients participated and 118 (81%) completed the PGI correctly. The ICC was 0.87, SEM 7.25 and SDC 20.10. Ninety-three per cent of the hypotheses of correlation were confirmed in tests for construct validity. Responsiveness was confirmed in 53% and 71% of hypotheses tested at 5 and 52 weeks. SRMs were 0.2 and 0.4, respectively.

Conclusion. The results support the validity, reliability and responsiveness of the Norwegian version of the PGI in patients with rheumatic diseases and its application as an outcome measure in rehabilitation or self-management programmes. Further research is needed to improve completion rates for the PGI.

Key words: Patient Generated Index, rheumatic diseases, rehabilitation, outcome measure, psychometrics.

CLINICAL SCIENCE

Introduction

Despite improvements in medical treatment, many patients with rheumatic diseases continue to be in need of rehabilitation [1]. Comprehensive rehabilitation is characterized by being tailored to patients' individual needs and challenges with the aims of reducing the consequences of illness and improving activity, participation and quality of life [2]. This presents challenges when selecting patient-reported outcome (PRO) measures for assessing the effectiveness of rehabilitation. Qualitative studies describe the beneficial effects of rehabilitation [2]; however, few quantitative studies have shown long-term effectiveness [3]. This may in part be due to the choice of PRO instruments, most of which are based on summated rating scales that are fixed or standardized in terms of content [4] and may have limited relevance to patients and their rehabilitation.

[1]National Resource Center for Rehabilitation in Rheumatology, Diakonhjemmet Hospital, [2]University of Oslo, Institute of Health and Society, [3]FORMI (Communication Unit for Musculoskeletal Disorders), Oslo University Hospital, Ullevaal and [4]Norwegian Knowledge Centre for the Health Services, Oslo, Norway.

Submitted 18 April 2012; revised version accepted 28 November 2012.

Correspondence to: Mari Klokkerud, National Resource Center for Rehabilitation in Rheumatology, Diakonhjemmet Hospital, PO Box 23, 0319 Oslo, Norway.
E-mail: mklokkerud@hotmail.com/mari.klokkerud@diakonsyk.no

It has been argued that only the individual is in a position to define his/her quality of life and that standardized instruments therefore may lack content validity in a patient perspective [4]. Individualized instruments have been developed, asking the patient to choose items or domains and weight their importance. The exclusion of aspects of life that are not of direct concern for the individual may reduce the noise that is present in standardized instruments, which in theory has the potential to make individualized instruments more responsive [5].

The Patient Generated Index (PGI) is an individualized instrument that is based on Calman's definition of quality of life: 'the extent to which our hopes and ambitions are matched by experience' [6]. Calman's description of the aim of all health care as 'helping the patient to narrow the gap between a patient's hopes and expectations and what actually happens' reflects the aims of rehabilitation. The PGI has been used since the early 1990s [7] and has been adapted and evaluated in various patient populations and languages [5]. The instrument is reliable for group comparisons and has shown good validity and responsiveness [5]. PGI has been used in patients with hip and knee OA [8], AS [9], elderly people with arthritis [10] and generalized musculoskeletal disorders [11]. It has also been recommended as an evaluative tool for AS [12]. The responsiveness of PGI has not been tested in patients with rheumatic diseases participating in rehabilitation or self-management programmes, neither has it been tested in a Norwegian population.

The objective of this study was to evaluate the psychometric properties of the Norwegian version of the PGI in patients with rheumatic diseases participating in rehabilitation or self-management programmes. The PGI was assessed for missing data, test–retest reliability, construct validity and responsiveness.

Patients and methods

Sample and data collection

Adults aged 18–75 years diagnosed with a rheumatic disease attending inpatient rehabilitation or self-management programmes between December 2008 and September 2009 were invited to participate in this study. The rehabilitation programmes lasted 2–4 weeks and were situated at two rehabilitation centres. The 1-week self-management programme was conducted at a hospital. None of the programmes had any standardized follow-up intervention. Patients with cognitive dysfunction or problems with reading or writing Norwegian were excluded.

The testing of reliability was limited to patients participating in the self-management programmes, who completed baseline questionnaires mailed to their homes 1–3 days before admission and on arrival. Patients at the rehabilitation centres completed the baseline questionnaire on arrival. A follow-up questionnaire was sent by mail to all participants approximately 5 and 52 weeks after inclusion, with reminders sent to non-responders after 3 weeks.

All patients received information about the study and gave written consent. The study was approved by the Norwegian Regional Committee for Medical Research Ethics and the Data Inspectorate.

Outcome measures

An English version of the PGI for AS [9] with improved scaling that has been used in different populations [13–15] was translated into Norwegian following recommended criteria [16]. Two translators fluent in English and Norwegian individually translated the PGI from English to Norwegian and any disagreement was resolved by consensus. The trigger list of areas was based on relevant content from existing lists, including the AS version [9]. Experienced clinicians working within rheumatology rehabilitation and eight patient research partners tested the Norwegian version and added identity and sexuality to the trigger list. The final version was successfully piloted in 20 patients with rheumatic diseases, and no changes were made. Two other translators fluent in English and Norwegian carried out back-translation. Disagreements were solved in a consensus meeting. The two English versions were found to be equivalent with no differences in meaning.

The PGI is completed in three stages. In Stage 1 the patient is asked to list up to five areas of life that are affected by his/her rheumatic condition. A trigger list of areas is provided that they may consult. A predefined sixth area is listed as 'all other areas of your life affected by your rheumatic disease'. In Stage 2 the patient scores the areas given in Stage 1 on a 7-point scale from as bad as could possibly be to as good as could possibly be. In Stage 3 the patient has 10 points to allocate to their areas according to where they would most value an improvement.

PGI scores are calculated by multiplying the area ratings in Stage 2 by the proportion of points given to the area in Stage 3, summing, dividing the sum by the number of areas described and finally multiplying by 10 to give a score between 0 and 100. Higher scores reflect better quality of life. A closed follow-up format of the PGI was used, where the areas defined by the patient at baseline were rated at all follow-up assessments.

In addition to the PGI, pain, fatigue and disease activity were assessed on a numerical rating scale (NRS) from 0 to 10, where 10 is the worst possible. Functional status was assessed by a Norwegian version of the arthritis-specific Modified Health Assessment Questionnaire (MHAQ), which comprises eight questions about the ability to perform specific activities, scored from 0 (no problem) to 3 (unable to do). The Norwegian version of the MHAQ has been widely applied [17–19], but there is no published evaluation of reliability and validity. General health was assessed by a Norwegian version of the Medical Outcomes Study Short Form 36-item Health Survey (SF-36), which has been evaluated in patients with rheumatic diseases [20]. The SF-36 items, which include descriptive rating scales, sum up to give eight scores representing different aspects of health on a scale from 0 to 100, where 100 is the best possible health. Two summary scores, the physical component score (PCS) and mental

component score (MCS), were calculated from the eight scale scores [21]. PCS and MCS use norm-based scoring from the general population and are transformed to have a mean of 50 (S.D. = 10).

Quality of life was assessed by the Norwegian version of the Quality of Life Scale (QoLS) [22]. Patients rate their level of satisfaction with 16 areas on a 7-point descriptive scale from very dissatisfied to very satisfied. The QoLS is scored by summing items to give a total score from 16 to 112, where 112 is the best possible quality of life [23]. The Norwegian QoLS has evidence for reliability and validity within patients with psoriasis [22]. It has not been evaluated in patients with rheumatic conditions but has been used in different patient populations [24–26] and a healthy Norwegian population [27].

Statistical analysis

The analysis followed the Consensus-based Standards for the Selection of Health Status Measurement Instruments (COSMIN) checklist [28] and the quality criteria proposed by Terwee *et al.* [29] for measurement properties of health status questionnaires.

Reliability

Test–retest reliability was assessed by the intraclass correlation coefficient (ICC) with the criterion of 0.7 and above for adequate reliability [29, 30]. Agreement was assessed by standard error of measurement (SEM) and smallest detectable change (SDC) for individual and group comparisons. The SEM is equal to the square root of the residuals of the mean square of an analysis of variance [29]. The SDC for individuals (SDCind) is equal to 1.96 multiplied by $\sqrt{2}$ and multiplied by the SEM The SDC for group comparisons (SDCgroup) was calculated by SDCind divided by $\sqrt{n}$ [29].

Validity

Construct validity has been defined as the degree to which instrument scores are consistent with hypotheses that are based on assumptions that the instrument validly measures the construct under consideration [31]. Construct validity was based on *a priori* hypotheses for levels of correlation (Pearson's *r*) between baseline PGI scores, which were normally distributed, and general health (SF-36), quality of life (QoLS), function and activity (MHAQ), symptoms and disease activity (NRS for pain, fatigue and disease activity). High correlation was defined as $r \geqslant 0.60$, moderate correlation as $0.30 \geqslant r < 0.60$ and low correlation as $r < 0.30$ [32]. In completing the PGI, patients include areas of personal importance as well as relative weights in the form of importance weightings. These two characteristics set the PGI apart from the standardized instruments included in the questionnaire. It is possible that two patients may have similar scores on the standardized instruments but very different PGI scores, hence high levels of correlation were not expected. Since rheumatic diseases primarily lead to reduced physical function and increased physical symptoms like pain, low levels of correlation were hypothesized for the SF-36 role emotional, mental health and MCS scales. For the remainder of the SF-36, including PCS and MHAQ, NRS for pain, fatigue and disease activity and QoLS, moderate levels of correlation were expected.

Responsiveness

Responsiveness has been defined as the ability of an instrument to detect changes over time in the construct to be measured [28]. The COSMIN recommend that an *a priori* hypothesis regarding the correlation between different change scores should be included in the evaluation of responsiveness [28]. Based on former studies, PGI and other individualized instruments [5, 33, 34], larger changes in PGI scores were expected than for the other instruments due to the individual format. Changes in PGI scores at 5 and 52 weeks were hypothesized as moderately correlated with changes in scores for the SF-36 PCS, QoLS and NRS for pain, fatigue and disease activity. Low correlations were hypothesized with changes in scores for the SF-36 MCS and MHAQ.

The standardized response mean (SRM) was used to compare the relative responsiveness of the different study instruments and is equal to the mean change score divided by the S.D. of the mean change [35]. For the statistical analysis, SPSS for Windows 17.0 was used.

Results

Patients and data quality

Of the 145 patients who agreed to participate, 90 participated in a self-management programme and 55 in a rehabilitation programme. There were no statistical differences between patients attending the two clinical settings with respect to health status and sociodemographic variables. However, patients at the rehabilitation centres had a statistically significant longer disease duration compared with patients participating in the self-management programmes [mean (± S.D.) 11.0 (8.4) *vs* 5.8 (6.2) years, $P < 0.001$). The PGI was correctly completed by 118 (81%) patients at baseline. Statistically significant longer disease duration was found for the non-completers, but there were no other differences between the two groups (Table 1).

Baseline PGI scores were normally distributed with a mean (± S.D.) of 39.3 (14.2). The 118 patients gave 576 areas that could be classified into 22 categories (Fig. 1). The five most frequently listed areas according to category were fatigue, pain, sleep, socializing and housework. Eighty-five per cent of all areas listed were chosen from the trigger list.

Reliability

The retest questionnaire was completed by 60 of the 70 invited patients (85.7%), of whom 41 completed both the test and retest PGI correctly. These patients had a mean (± S.D.) baseline PGI score of 38.05 (14.18), which was slightly lower but not statistically significant from the total population score of 39.28 (14.19). The mean (± S.D.) retest score was 41.10 (15.27), which was not significantly different from the baseline scores. The SEM was 7.25,

Table 1 Baseline characteristics of patients completing and not completing the patient-generated index

	Total respondents ($n = 145$)	PGI completed ($n = 118$)	PGI not completed ($n = 27$)
Age, mean (S.D.)	54.6 (10.5)	54.1 (10.1)	56.9 (11.8)
Gender, *n* (%)			
Male	18 (12.4)	14 (11.9)	4 (14.8)
Female	127 (87.6)	104 (88.1)	23 (85.2)
Education >12 years, *n* (%)	57 (39.3)	50 (42.4)	7 (25.9)
Diagnosis, *n* (%)			
RA	16 (11.1)	12 (10.3)	4 (14.8)
AS	7 (4.9)	5 (4.3)	2 (7.4)
CTD	26 (18.1)	22 (18.8)	4 (14.8)
PMR	1 (0.7)	1 (0.9)	0
OA	25 (17.4)	20 (17.1)	5 (18.5)
FM	18 (12.5)	14 (12.0)	4 (14.8)
Other rheumatic diseases	1 (0.7)	1 (0.9)	0
More than one rheumatic disease	50 (34.7)	42 (35.9)	8 (29.6)
Missing	1 (0.7)	1 (0.9)	0
Years since diagnosis, mean (S.D.)	7.8 (7.5)	7.5 (7.8)	10.9 (10.3)*
Comorbidities, mean (S.D.)	2.2 (1.6)	2.2 (1.6)	2.3 (1.4)
PGI, mean (S.D.)		39.3 (14.2)	
MHAQ,[a] mean (S.D.)	0.5 (0.4)	0.5 (0.4)	0.5 (0.4)
NRS,[b] mean (S.D.)			
Pain	5.6 (2.0)	5.6 (1.9)	5.4 (2.5)
Fatigue	6.6 (2.3)	6.7 (2.3)	6.4 (2.5)
Disease activity	5.7 (2.0)	5.7 (1.9)	5.7 (2.2)
SF-36,[c] mean (S.D.)			
Physical functioning	56.3 (19.9)	56.9 (19.4)	53.33 (22.2)
Social functioning	57.9 (26.4)	57.4 (27.0)	60.19 (23.8)
Role physical	43.1 (24.2)	42.7 (23.6)	45.25 (27.3)
Role emotional	68.0 (30.2)	70.1 (29.8)	58.0 (30.9)
Mental health	67.8 (18.8)	68.4 (19.8)	65.4 (13.7)
Vitality	27.8 (17.4)	28.1 (17.7)	26.6 (16.3)
Pain	36.3 (18.2)	35.6 (16.2)	39.4 (25.2)
General health	41.0 (20.1)	41.4 (20.4)	39.2 (18.7)
PCS	33.7 (8.0)	33.6 (8.1)	34.6 (7.8)
MCS	43.4 (11.9)	43.9 (12.3)	41.0 (9.5)
QoLS,[d] mean (S.D.)	75.2 (13.5)	75.5 (14.0)	74.7 (11.7)

[a]The MHAQ assesses functional status and is scored from 0 to 3; 3 is the worst overall disability. [b]NRSs are scored from 0 to 10; 10 is the worst pain, fatigue and symptoms. [c]SF-36 scales are scored from 0 to 100; 100 is the best possible health. PCS and MCS use norm-based scoring from the general population and are transformed to have a mean of 50 (S.D. = 10). [d]The QoLS assesses experience of quality of life according to predefined areas and is scored from 16 to 112, where a higher score means better quality of life. *$P < 0.05$.

with an SDCind of 20.10 and an SDCgroup of 3.14. The ICC (95% CI) was 0.87 (0.74, 0.93).

Validity

The levels of correlation between the PGI baseline scores and the other instruments ranged from 0.23 to 0.47 (Table 2). Fourteen (93%) of the 15 hypotheses were confirmed, showing that the PGI correlated more strongly with the physical and disease-related symptoms than the mental and emotional aspects of the SF-36.

Responsiveness

All instruments were correctly completed by 80 and 69 patients after 5 and 52 weeks, respectively. There were no significant differences between those who filled in the PGI correctly and those who did not in change scores for the other instruments.

The responsiveness hypotheses were only partially supported. At 5 weeks, 53% of the hypotheses were supported and at 52 weeks, 71% were supported. As expected, changes in PGI scores were moderately correlated to changes in the QoLS at 5 weeks and the NRS of pain, fatigue and disease activity at 5 and 52 weeks (Table 2). However, contrary to what was expected, changes in PGI scores had a low correlation with those for the SF-36 PCS at both 5 and 52 weeks. Furthermore, the expected low correlation with changes in the MHAQ at 5 and 52 weeks were only confirmed at 52 weeks, the correlation at 5 weeks being moderate. The correlation with change scores for the SF-36 MCS were low at

FIG. 1 The 576 areas reported by 118 patients as affected by their rheumatic disease (categorized into 22 areas).

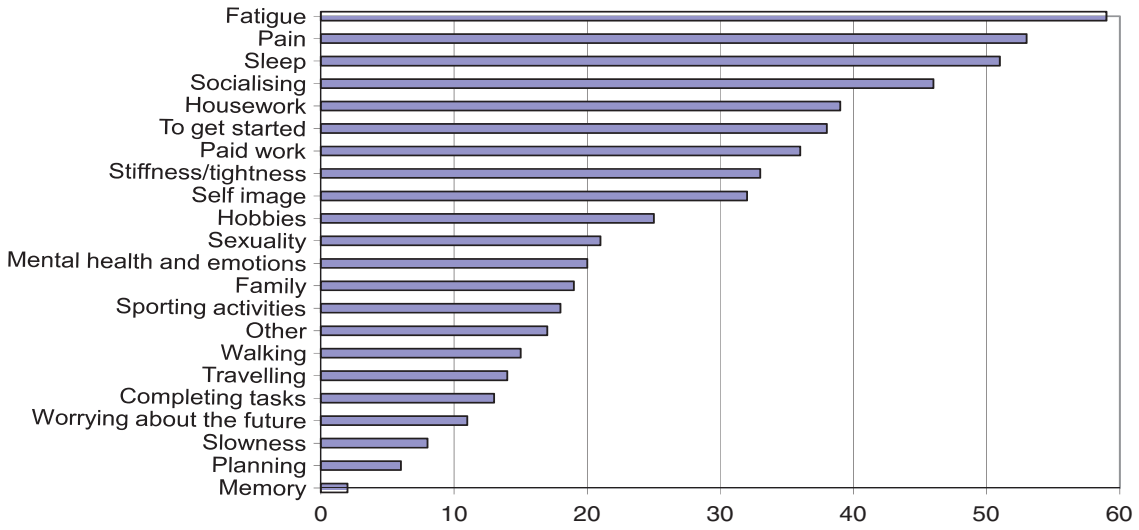

TABLE 2 Pearson correlations between the PGI and scores for MHAQ, NRS pain, fatigue and disease activity, SF-36 and QoLS

	Hypothesized correlation baseline and change	PGI baseline r ($n = 118$)	Change in PGI after 5 weeks r ($n = 80$)	Change in PGI after 52 weeks r ($n = 68$)
MHAQ[a]	Low	−0.26** S	0.35*	0.11 S
NRS[b]				
Pain	Moderate	−0.46** S	0.42** S	0.34* S
Fatigue	Moderate	−0.31** S	0.31* S	0.44**S
Disease activity	Moderate	−0.47** S	0.35* S	0.43** S
SF-36[c]				
Physical functioning	Moderate	0.35** S	0.25*	0.28*
Social functioning	Moderate	0.45** S	0.21*	0.32**S
Role physical	Moderate	0.47** S	0.24*	0.39** S
Role emotional	Low	0.23* S	0.23* S	0.23* S
Mental health	Low	0.27** S	0.25* S	0.50**
Vitality	Moderate	0.35** S	0.40** S	0.45** S
Pain	Moderate	0. 38** S	0.21	0.47** S
General health	Moderate	0.34** S	0.17	0.15 S
PCS	Moderate	0.39** S	0.20	0.29*
MCS	Low	0.30**	0.29* S	0.40*
QoLS[d]	Moderate	0.39** S	0.35* S	–

$r \geqslant 0.60$: high correlation; $0.30 \geqslant r < 0.60$: moderate correlation; $r < 0.30$: low correlation [32]. MHAQ and the three NRSs have the opposite scale to PGI. [a]The MHAQ assesses functional status and is scored from 0 to 3; 3 is the worst overall disability. [b]NRSs are scored from 0 to 10; 10 is the worst pain, fatigue and symptoms. [c]SF-36 scales are scored from 0 to 100; 100 is the best possible health. PCS and MCS use norm-based scoring from the general population and are transformed to have a mean of 50 (S.D. = 10). [d]The QoLS assesses experience of quality of life according to predefined areas and is scored from 16 to 112, where a higher score means better quality of life. *$P < 0.05$; **$P < 0.001$. S: hypothesis supported.

5 weeks but moderate at 52 weeks, being higher than expected.

The change scores and SRMs are presented in Table 3. The PGI, NRS pain and SF-36 scales of social functioning, role physical and vitality showed significant improvement at 5-week follow-up. The mean PGI change of 3.9 (17.0) and 6.4 (15.4) at 5 and 52 weeks, respectively, was higher than the SDCgroup, but lower than the SDCind. The SF-36 scale of vitality had the highest SRM of 0.4 after 5 weeks. One year after rehabilitation, the PGI had the highest SRM of 0.4.

TABLE 3 Mean change (95% CI) and standard response mean (SRM) (95% CI) 5 and 52 weeks after rehabilitation or self-management programmes

	Admission 1[a]	5 weeks	SRM 5 weeks	Admission 2[a]	52 weeks	SRM 52 weeks
PGI	39.4 (36.2, 42.7)	42.7 (39.7, 45.8)	0.2 (0.0, 0.5)	38.6 (35.1, 42.0)	42.9 (39.3, 46.5)	0.4 (0.2, 0.7)
MHAQ[b]	0.5 (0.4, 0.6)	0.5 (0.4, 0.6)	0.1 (−0.2, 0.3)	0.4 (0.4, 0.6)	0.5 (0.4, 0.6)	0.1 (−0.1, 0.4)
NRS[c]						
Pain	5.5 (5.1, 5.9)	5.2 (4.9, 5.6)	0.1 (0.1, 0.3)	5.5 (5.0, 6.0)	5.4 (5.0, 5.8)	0.3 (0.1, 0.5)
Fatigue	6.7 (6.2, 7.2)	6.2 (5.8, 6.7)	0.2 (−0.1, 0.4)	6.8 (6.3, 7.3)	6.6 (6.1, 7.1)	0.1 (−0.2, 0.3)
Disease activity	5.7 (5.3, 6.2)	5.3 (5.0, 5.7)	0.2 (−0.1, 0.4)	5.8 (5.3, 6.2)	5.5 (5.1, 5.9)	0.1 (−0.1, 0.4)
SF-36[d]						
Physical function	57.0 (52.4, 61.6)	57.7 (54.1, 61.3)	0.2 (−0.0, 0.4)	57.2 (52.5, 61.9)	58.0 (53.6, 62.5)	0.2 (−0.1, 0.4)
Social function	59.2 (53.2.65.2)	63.1 (58.5, 67.8)	0.2 (0.0, 0.5)	55.0 (48.8, 61.1)	54.3 (48.7, 59.8)	−0.1 (−0.3, 0.2)
Role physical	43.4 (38.0, 48.7)	49.3 (44.8, 53.8)	0.3 (0.1, 0.5)	42.2 (36.9, 47.5)	47.9 (42.7, 53.2)	0.2 (−0.0, 0.5)
Role emotional	71.8 (65.1, 78.4)	70.8 (65.6, 76.0)	0.1 (−0.1, 0.3)	71.3 (64.4, 78.2)	70.4 (65.0, 75.8)	0.1 (−0.1, 0.4)
Mental health	69.6 (65.5, 73.8)	67.4 (63.6, 71.1)	0.1 (−0.1, 0.3)	68.5 (64.1, 73.0)	67.3 (63.3, 71.2)	0.0 (−0.2, 0.3)
Vitality	28.5 (24.3, 32.7)	34.8 (31.2, 38.4)	0.4 (0.1, 0.6)	26.0 (21.8, 30.2)	32.3 (28.2, 36.4)	0.3 (0.1, 0.6)
Pain	36.7 (33.2, 40.1)	39.0 (35.7, 42.3)	0.1 (−0.1, 0.3)	36.2 (32.8, 39.5)	39.0 (35.2, 42.8)	0.1 (−0.1, 0.3)
General health	44.3 (40.1, 48.5)	44.1 (40.7, 47.6)	0.1 (−0.2, 0.3)	42.3 (38.0, 46.6)	42.6 (38.7, 47.3)	0.1 (−0.1, 0.3)
PCS	33.7 (31.8, 35.6)	35.4 (33.9, 36.8)	0.3 (0.0, 0.5)	33.4 (31.5, 35.4)	35.1 (33.2, 36.9)	0.2 (−0.1, 0.4)
MCS	44.7 (42.0, 47.4)	44.6 (42.3, 46.9)	0.2 (−0.1, 0.4)	43.6 (40.7, 46.5)	43.4 (41.0, 45.7)	0.1 (−0.2, 0.3)
QoLS[e]	77.0 (74.2, 79.8)	75.3 (72.5, 78.0)	0.1 (−0.1, 0.2)	76.1 (72.8, 79.3)	NR	NR

Positive changes represent improvement. [a]The two sets of admission scores are for patients completing all instruments at baseline and 5 weeks (Admission 1) and at baseline and 52 weeks (Admission 2). [b]The MHAQ assesses functional status and is scored from 0 to 3; 3 is the worst overall disability. [c]NRSs are scored from 0 to 10; 10 is the worst pain, fatigue and symptoms. [d]SF-36 scales are scored from 0 to 100; 100 is the best possible health. PCS and MCS use norm-based scoring from the general population and are transformed to have a mean of 50 (S.D. = 10). [e]The QoLS assesses experience of quality of life according to predefined areas and is scored from 16 to 112, where a higher score means better quality of life.

Discussion

The results of this study support the reliability, validity and responsiveness of the Norwegian version of the PGI in patients with rheumatic diseases and its application as an outcome measure in rehabilitation or self-management programmes. These properties are necessary for an instrument designed for assessing PROs, and the COSMIN recommendations have been followed [36].

Reliability

The ICC was above the recommended level of 0.70, which suggests that the PGI is a reliable outcome measure in this patient group. Similar levels of reliability have been found in other studies with comparable versions and populations [9, 37]. The PGI had a relatively high measurement error of 20.10 and 3.14 at the individual and group levels, respectively. This is slightly higher than what was previously reported for the SF-36 summary scales in a comparable population, with an SDCind of 12.0 and SDCgroup of 1.39 for the PCS and an SDCind of 16.0 and SDCgroup of 1.94 for the MCS [38]. As far as we know, SDC has not been reported in previous evaluations of the PGI.

Validity

As expected, the correlation with scores for the other instruments was low to moderate, and 93% of the hypothesis was supported, indicating good construct validity. Two previous studies within rheumatology gave slightly different findings. The PGI scores were found to have moderate correlation with those for both the SF-12 summary scales in UK patients with AS [9]. The current study found low correlation with the SF-36 MCS, which was as expected. In a previous study of disabled patients, the PGI scores correlated most strongly with the SF-36 scales of social functioning (0.49) and vitality (0.47) [39]. The current study found a moderate correlation with the SF-36 social functioning scale that was among the highest, and that for the vitality scale was relatively lower. The different levels of correlation with a generic instrument such as the SF-36 in these study populations might reflect the different characteristics of the patient populations, which influences choice of areas, scoring and points given to areas in PGI completion.

The PGI is based on the assumption that health problems affect individuals and their quality of life differently and therefore are best defined by the individual patient. Giving patients the opportunity to choose what is measured minimizes threats to the content validity [40]. As in other studies [9, 10], most patients chose areas that were included in the trigger list. Patient research partners initially tested the instrument (including the trigger list), which strengthens the content validity of the PGI. In completing the PGI, patients nominate their own areas, score them and allocate points in terms of their relative importance, thereby generating an individualized item pool with a form of weighting attached. This makes the assessment of internal construct validity using widely

applied methods such as factor analysis and modern psychometric methods such as Rasch analysis problematic. However, in a larger study, validity was assessed by comparing the PGI scores for subgroups of patients whose areas largely related to specific aspects of health, such as physical function, with the corresponding scale scores for the SF-36 [7]. This method is recommended for future studies with larger sample sizes.

Responsiveness

According to the SRMs, the PGI was among the most responsive instruments at 5-week follow-up, with only the SF-36 scales of role physical and vitality having SRMs >0.30. As hypothesized, PGI change scores at 5 weeks had low to moderate correlation with the other PROs used, with the highest correlation with pain and SF-36 vitality. One year later the PGI was found to be the most responsive instrument with the highest SRM. As expected, the other instrument scores were low to moderately correlated with those for the PGI. Contrary to expectations, changes in scores for the SF-36 mental health scale had the highest correlation with those for the PGI. Future research is recommended for assessing the relationship between PGI scores, mental health and other aspects of health over time. The mean PGI change both at 5 and 52 weeks was higher than the SDCgroup, which strengthens the belief of a true and significant change. It has been suggested that patients may be more likely to register changes in areas they regard as important than in areas they consider less important or are indifferent towards [40]. The individualized nature of the items in the PGI may therefore increase the likelihood that changes are registered and valued as significant by the patients.

For satisfactory outcomes, the rehabilitation process should be closely linked to behavioural changes, in which time is an important component [2]. To adapt and implement new strategies in everyday life after rehabilitation takes time, which might explain the higher SRM of the PGI at 52 weeks compared with 5 weeks after rehabilitation.

Most standardized instruments have been developed to assess PROs following standardized interventions [4]. The diversity of goals in rehabilitation, along with the individual adaptation of interventions, makes it challenging to use standardized instruments that include items based on aspects of health and quality of life that have general relevance to a specific patient population but might not be relevant to the individual patient.

Acceptability and data quality

Compared with instruments with standardized items, individualized instruments may be more burdensome to complete. The PGI is one of few individualized instruments that largely have been self-reported through postal surveys. The completion rate for the PGI in this study was higher than those reported previously [7, 10, 34, 39]. The reason for this may be the choice of a shorter scale in Stage 2 and the subsequent need for fewer points, in this case 10, in Stage 3. A structured review found higher levels of reliability for versions with shorter scales [5]. An early study found that patients who completed the PGI were on average younger, better educated and had higher incomes than the non-completers [41]. In the current study, completers had shorter disease duration compared with non-completers. Long disease duration may indicate poorer health and thus less cognitive energy to understand instructions and complete an instrument. However, no statistically significant differences between the completers and non-completers were found for other health variables. Completion rates may be improved through improvements in written instructions and assistance during completion. Future electronic or online versions may also improve acceptability and completion rates.

Strengths and limitations

The diagnostic heterogeneity and high prevalence of comorbidity among participants is a study strength. This is common within rehabilitation in rheumatology [42] and hence is important for generalizability.

It can be questioned whether the interval between the test and the retest of reliability was long enough to ensure that participants could not recall previous responses. However, considering the fluctuating nature of these diseases, a longer interval could jeopardize the demanded stability of disease activity. Moreover, the nature of completion for the PGI, with its three stages, will reduce recall as compared with standardized instruments.

The first follow-up in this study was after 5 weeks, which is relatively short for measuring changes after rehabilitation interventions. An additional follow-up at 6 months would have strengthened our study and provided important information about changes over time.

Conclusion

In conclusion, this study shows that the Norwegian version of the PGI is a reliable, valid and responsive instrument for measuring individualized quality of life in patients with rheumatic diseases. In comparison with standardized instruments based on summed rating scales, the individualized nature of the PGI makes it appropriate for use in the long-term evaluation of rehabilitation and self-management programmes. To increase the acceptability of the PGI, electronic versions with corrective feedback or the presence of assistance at baseline should be considered.

Rheumatology key messages

- The Norwegian version of the PGI is valid and reliable in patients with rheumatic diseases.
- Individualized instruments seem particularly appropriate for evaluating rehabilitation and self-management programmes in rheumatic diseases.
- Future electronic versions of the PGI may improve completion rates in rheumatic diseases.

Acknowledgements

We thank Turid Dager, specialist in psychology at Diakonhjemmet Hospital, and the NRRK Patient Panel for their important contributions to pretesting the PGI. We also thank Marit Tangerud, Lillehammer Hospital for Rheumatic Diseases, Ida Korsvold and Gilles Jarret, Skogli Rehabilitation Center, and Kjell Haavik-Nilsen and Jon Arthur Innerdahl, Vikersund Rehabilitation Center, for essential help with data collection.

Disclosure statement: The authors have declared no conflicts of interest.

References

1 Petersson IF. Evolution of team care and evaluation of effectiveness. Curr Opin Rheumatol 2005;17:160–3.

2 Dager TN, Kjeken I, Fjerstad E *et al*. 'It is about taking grips and not let myself be ravaged by my body': a qualitative study of outcomes from in-patient multidisciplinary rehabilitation for patients with chronic rheumatic diseases. Disabil Rehabil 2012;34:910–6.

3 Vliet Vlieland TPM. New models of care for patients with rheumatoid arthritis. Expert Rev Pharmacoecon Outcomes Res 2006;6:159–69.

4 Garratt A, Schmidt L, Mackintosh A *et al*. Quality of life measurement: bibliographic study of patient assessed health outcome measures. BMJ 2002;324:1417.

5 Martin F, Camfield L, Rodham K *et al*. Twelve years' experience with the Patient Generated Index (PGI) of quality of life: a graded structured review. Qual Life Res 2007;16:705–15.

6 Calman KC. Quality of life in cancer patients—an hypothesis. J Med Ethics 1984;10:124–7.

7 Ruta DA, Garratt AM, Leng M *et al*. A new approach to the measurement of quality of life: the Patient-Generated Index. Med Care 1994;32:1109–26.

8 Hawker GA, Stewart L, French MR *et al*. Understanding the pain experience in hip and knee osteoarthritis—an OARSI/OMERACT initiative. Osteoarthritis Cartilage 2008; 16:415–22.

9 Haywood KL, Garratt AM, Dziedzic K *et al*. Patient centered assessment of ankylosing spondylitis-specific health related quality of life: evaluation of the Patient Generated Index. J Rheumatol 2003;30:764–73.

10 Tully MP, Cantrill JA. The validity of the modified patient generated index—a quantitative and qualitative approach. Qual Life Res 2000;9:509–20.

11 Jolles BM, Buchbinder R, Beaton DE. A study compared nine patient-specific indices for musculoskeletal disorders. J Clin Epidemiol 2005;58:791–801.

12 van der Heijde D. New directions in classification and outcome assessment in ankylosing spondylitis. Curr Rheumatol Rep 2004;6:98–101.

13 Camfield L, Ruta D. 'Translation is not enough': using the Global Person Generated Index (GPGI) to assess individual quality of life in Bangladesh, Thailand, and Ethiopia. Qual Life Res 2007;16:1039–51.

14 Tavernier SS, Beck SL, Clayton MF *et al*. Validity of the Patient Generated Index as a quality-of-life measure in radiation oncology. Oncol Nurs Forum 2011; 38:319–29.

15 Tavernier SS, Totten AM, Beck SL. Assessing content validity of the patient generated index using cognitive interviews. Qual Health Res 2011;21:1729–38.

16 Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993;46:1417–32.

17 Lie E, van der Heijde D, Uhlig T *et al*. Effectiveness and retention rates of methotrexate in psoriatic arthritis in comparison with methotrexate-treated patients with rheumatoid arthritis. Ann Rheum Dis 2010;69:671–6.

18 Uhlig T, Heiberg T, Mowinckel P *et al*. Rheumatoid arthritis is milder in the new millennium: health status in patients with rheumatoid arthritis 1994–2004. Ann Rheum Dis 2008;67:1710–5.

19 Uhlig T, Moe R, Reinsberg S *et al*. Responsiveness of the International Classification of Functioning, Disability and Health (ICF) Core Set for rheumatoid arthritis. Ann Rheum Dis 2009;68:879–84.

20 Loge JH, Kaasa S, Hjermstad MJ *et al*. Translation and performance of the Norwegian SF-36 Health Survey in patients with rheumatoid arthritis. I. Data quality, scaling assumptions, reliability, and construct validity. J Clin Epidemiol 1998;51:1069–76.

21 Ware JE. SF-36® Health Survey Update. SF-36 org www.sf-36.org/tools/sf36.shtml (18 March 2012, date last accessed).

22 Wahl A, Burckhardt C, Wiklund I *et al*. The Norwegian version of the Quality of Life Scale (QOLS-N). A validation and reliability study in patients suffering from psoriasis. Scand J Caring Sci 1998;12:215–22.

23 Burckhardt CS, Woods SL, Schultz AA *et al*. Quality of life of adults with chronic illness: a psychometric study. Res Nurs Health 1989;12:347–54.

24 Wahl AK, Rustoen T, Hanestad BR *et al*. Living with cystic fibrosis: impact on global quality of life. Heart Lung 2005; 34:324–31.

25 Rohde G, Haugeberg G, Mengshoel AM *et al*. Is global quality of life reduced before fracture in patients with low-energy wrist or hip fracture? A comparison with matched controls. Health Qual Life Outcomes 2008;6:90.

26 Moi AL, Wentzel-Larsen T, Salemark L *et al*. Impaired generic health status but perception of good quality of life in survivors of burn injury. J Trauma 2006;61:961–8.

27 Wahl AK, Rustoen T, Hanestad BR *et al*. Quality of life in the general Norwegian population, measured by the Quality of Life Scale (QOLS-N). Qual Life Res 2004;13: 1001–9.

28 Mokkink LB, Terwee CB, Knol DL *et al*. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC Med Res Methodol 2010;10:22.

29 Terwee CB, Bot SD, de Boer MR *et al*. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34–42.

30 Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. 3rd edn. New York: Oxford University Press, 2003.

31 Mokkink LB, Terwee CB, Patrick DL *et al*. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010;63:737–45.

32 Andresen EM. Criteria for assessing the tools of disability outcomes research. Arch Phys Med Rehabil 2000;81: S15–20.

33 Kjeken I, Slatkowsky-Christensen B *et al*. Norwegian version of the Canadian Occupational Performance Measure in patients with hand osteoarthritis: validity, responsiveness, and feasibility. Arthritis Rheum 2004;51:709–15.

34 Ruta DA, Garratt AM, Russell IT. Patient centred assessment of quality of life for patients with four common conditions. Qual Health Care 1999;8:22–9.

35 Terwee CB, Dekker FW, Wiersinga WM *et al*. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res 2003;12:349–62.

36 Mokkink LB, Terwee CB, Patrick DL *et al*. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010;19:539–49.

37 Witham MD, Fulton RL, Wilson L *et al*. Validation of an individualised quality of life measure in older day hospital patients. Health Qual Life Outcomes 2008;6:27.

38 Hurst NP, Ruta DA, Kind P. Comparison of the MOS short form-12 (SF12) health status questionnaire with the SF36 in patients with rheumatoid arthritis. Br J Rheumatol 1998; 37:862–9.

39 Macduff C, Russell E. The problem of measuring change in individual health-related quality of life by postal questionnaire: use of the patient-generated index in a disabled population. Qual Life Res 1998;7:761–9.

40 Kloppenburg M, Stamm T, Watt I *et al*. Research in hand osteoarthritis: time for reappraisal and demand for new strategies. An opinion paper. Ann Rheum Dis 2007;66: 1157–61.

41 Ternent L, Vale L, Buckley B *et al*. Measuring outcomes of importance to women with stress urinary incontinence. BJOG 2009;116:719–25.

42 Klokkerud M, Hagen KB, Lochting I *et al*. Does the content really matter? A study comparing structure, process, and outcome of team rehabilitation for patients with inflammatory arthritis in two different clinical settings. Scand J Rheumatol 2012;41:20–8.