RHEUMATOLOGY

Rheumatology 2014;53:678–685
doi:10.1093/rheumatology/ket423
Advance Access publication 18 December 2013

Original article

The Gout Assessment Questionnaire 2.0: cross-cultural translation into Dutch, aspects of validity and linking to the International Classification of Functioning, Disability and Health

Bart Spaetgens[1], Sjef van der Linden[2,3] and Annelies Boonen[2,3]

Abstract

Objectives. The Gout Assessment Questionnaire 2.0 (GAQ2.0) is a disease-specific patient-reported outcome measure for gout that distinguishes five different subscales and comprises overall 31 questions. The aims of this study were to translate the GAQ2.0 into Dutch and to test clinimetric properties.

Methods. Recommendations for translation and cross-cultural adaptation were followed and no cultural adaptations were needed. The resulting Dutch GAQ2.0 was administered to patients registered at the rheumatology outpatient clinic diagnosed with gout. Internal consistency was tested using Cronbach's α, reliability using intraclass correlation coefficient (ICC), content validity by linkage to the International Classification of Functioning, Disability and Health (ICF) and construct validity by correlating the subscales of the GAQ2.0 with the HAQ disability index (HAQ-DI) and 36-item Short Form Health Survey (SF-36).

Results. A total of 126 patients [106 (84%) male, mean age 66.6 years (S.D. 10.4), mean disease duration 11.2 years (S.D. 10.6)] completed a number of questionnaires, including the GAQ2.0, HAQ-DI and SF-36, and underwent a clinical examination. Internal consistency was sufficient (Cronbach's $\alpha = 0.83$–0.94), except for the subscale gout medication side effects (Cronbach's $\alpha = 0.51$). Test–retest reliability was good (ICCs 0.73–0.86) for all subscales, but moderate for the subscale unmet gout treatment need (ICC 0.56). Gout impact (GI) subscale scores showed only weak to moderate correlations with HAQ-DI and SF-36, but stronger emphasis on the emotional consequences of gout. Also, it correlated better with gout-specific outcomes such as the number of gout flares and pain.

Conclusion. The Dutch GAQ2.0 shows sufficient evidence of validity to assess disease-specific functioning and health in patients with gout and seems to capture different aspects than those represented in the HAQ and SF-36.

Key words: gout, Gout Assessment Questionnaire, quality of life, validation, patient reported outcome, validity.

CLINICAL SCIENCE

Introduction

Gout is a chronic inflammatory rheumatic disease caused by the precipitation of uric acid crystals in the synovial fluid and tissues [1]. Classically the disease is characterized by acute and transient arthritis of one or more joints, alternating with symptom-free episodes between gout attacks. A subgroup of patients develops frequent and prolonged attacks, and even chronic arthritis can develop. This may be associated with so-called tophi, which represent the accumulation of monosodium urate crystals in and around the joints. Both tophi and frequent gout flares or chronic arthritis can be associated with damage of cartilage, bone, skin and more rarely, other organs [2].

It is clear that the disease, with its unpredictable attacks characterized by severe pain and limitations in mobility and possible chronic discomfort due to joint damage,

[1]Department of Internal Medicine, Maastricht University Medical Center, [2]Department of Internal Medicine, Division of Rheumatology, Maastricht University Medical Center and [3]CAPHRI Research Institute, University Maastricht, Maastricht, The Netherlands.

Submitted 5 August 2013; revised version accepted 30 October 2013.

Correspondence to: Bart Spaetgens, Maastricht University Medical Center, PO Box 5800, NL-6202 AZ Maastricht, The Netherlands.
E-mail: bartholomeus.spaetgens@mumc.nl

chronic arthritis or tophi, can affect many aspects of health-related quality of life (HRQOL) [3–5]. Measuring the impact of gout on HRQOL from a patient's perspective is challenging because of the heterogeneous manifestations of gout that need to be captured. While generic instruments to assess (aspects of) HRQOL might be useful, it is expected that disease-specific questionnaires have a better content validity and therefore discriminative capacity, including sensitivity to change [6].

Colwell *et al.* [7] developed the first disease-specific Gout Assessment Questionnaire (GAQ) to assess the impact of gout across a broad range of areas relevant for patients' health and to be used in the setting of clinical trials. Candidate items were derived from a literature search, expert opinion and a limited number of patient interviews [8]. A second version of the Gout Assessment Questionnaire (GAQ2.0) was proposed by Hirsch *et al.* [9], after adjusting and optimizing the initial instrument for use in clinical practice. The GAQ2.0 showed acceptable reliability and validity in a community-based patient population and correlated more closely with patient-reported outcome measures of gout (e.g. frequent gout flares) than with the 36-item Short Form Health Survey (SF-36). In a later study, the discriminative ability of the GAQ2.0 according to the severity of gout was explored, which confirmed high scores correlate with more frequent gout attacks and more pain between gout attacks [10].

As such, the GAQ2.0 is the only patient-reported outcome (PRO) measure to assess the impact of gout on functioning and health. However, its application is limited by the existence of an English version only. The availability of versions in other languages could help to fill the gap with respect to a universally accepted disease-specific PRO to assess HRQOL in gout patients.

The purpose of the present study was to develop a Dutch version of the GAQ2.0 by performing a translation according to recommended methods and to assess further clinimetric properties of the Gout Impact Scale (GIS), which is the first section of the GAQ2.0 and represents typical areas of HRQOL that can be affected by gout. Specifically, internal consistency, test–retest reliability, content validity (comparison) with the International Classification of Functioning, Disability and Health (ICF), construct validity with the SF-36 and HAQ and discrimination between groups of gout patients with different disease severity of the GIS were assessed.

Materials and methods

Translation

The translation procedure was performed in collaboration with a reputable company (PharmaQuest, Banbury, UK) in translations of PROs. The translation, review, linguistic validation and cross-cultural adaptation process were performed according to the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) principles of good practice [11] and is consistent with the approach proposed by Beaton as best practice in rheumatology [12]. The forward translation was performed by two qualified native Dutch speakers who are experienced medical translators and speak fluent English. Back-translation of the Dutch version was performed by two qualified native English speakers who are experienced medical translators and speak fluent Dutch. After this, contact was sought with the developers of the original questionnaire, who approved the preceding steps as well as the final version. Finally, this version was pretested according to the concept of pilot testing and cognitive debriefing. Five patients under rheumatological care in our department were invited to participate. Attention was given to ensure they were native Dutch speakers and represented the spectrum of gender, age (three men, two women, age range 39–78 years), disease (disease duration range and acute intermittent or chronic gout) and education. Participants were first invited to complete the questionnaire. After completion, they were asked to read aloud the instructions, questions and scoring modalities line by line. After each line, they were asked to indicate whether the instruction/item/anchor was clear, were invited to repeat the sentences in their own words and were asked to think aloud while answering the questions. They were probed specifically to comment on whether sentences were unambiguous and whether they would have preferred another wording. Finally, they were asked whether the questions missed aspects of the disease that influence their functioning. These findings were listed and were returned to the translators, who discussed the need for adaptation. Ultimately, no cultural adaptations were necessary. Only minor adjustments in wording, grammar or typography were made. The patients indicated no aspects of their disease that they were missing. Time to complete the questionnaire ranged from 4 to 7 min. The final Dutch version is available as supplementary material, available at *Rheumatology* Online.

Assessment of clinimetric properties

Patients

A convenience sample of 250 patients that were registered with gout, according to the rheumatologist, in the diagnostic/administrative database of the department of rheumatology between January 2011 and April 2012 were invited by a letter to participate in the study. Patients that signed informed consent received an appointment about 4 weeks later. A random subsample of patients was asked again to complete the GAQ2.0 together with some questions on the recent course of their gout. The principles of the Declaration of Helsinki were followed and the study was approved by the ethics committee of Maastricht University Medical Center. Prior to the data collection, all participating patients signed the informed consent document.

Assessments

During the study visit, patients were interviewed about demographic characteristics (age, sex, education) and course of disease (symptom and disease duration, number of gout flares last year). Next, they underwent a clinical examination to determine the presence and

amount of tophi. Finally, patients completed a series of questionnaires.

Gout Assessment Questionnaire 2.0

The GAQ2.0 is a self-administered questionnaire consisting of 31 questions divided in two sections. First, the so-called GIS evaluates the current impact of gout in five areas: gout concern overall (4 items), gout medication side effects (2 items), unmet gout treatment needs (3 items), well-being during attack (11 items) and gout concern during attack (4 items). All subscales of the GISs are scored separately on a 0 to 100 score, with higher scores indicating a more important impact. The second section asks patients to describe whether they had a gout flare (yes/no) in the last year and to describe on a 6-point Likert scale to what extent the gout affected their quality of life, physical and mental health and pain in the past 4 weeks (1 = very poor, 6 = excellent). They were also asked to describe on a 10-point numeric rating scale how much pain they experienced due to gout (1 = no pain, 10 = severe pain) and disease activity (1= no disease activity, 10 = severe disease activity). Assessment of the clinimetric properties was limited to the 24 items of the GIS.

Health Assessment Questionnaire

The HAQ is an instrument to assess impairments in physical function in the last 7 days. It was developed for use in RA [13], but has been shown to be valid for use in other rheumatic diseases as well, such as gout [14]. The HAQ consists of 20 items across eight categories (dressing and grooming, arising, eating, walking, personal hygiene, reaching, gripping and other activities), scored on a 0- to 3-point Likert scale (0 = no difficulty, 3 = proposed action cannot be performed without help). The highest score per category is used and divided by 8, resulting in the HAQ-disability index (HAQ-DI) from 0 to 3 (higher score indicates worse physical functioning). A validated Dutch version of the HAQ is available [15].

The 36-Item Short Form Health Survey

The SF-36 is a generic instrument to assess HRQOL over the last 4 weeks. It consists of 36 questions in eight different domains: physical functioning (10 items), role limitation due to physical problems (4 items), bodily pain (2 items), general health perception (5 items), vitality (4 items), social functioning (2 items), role limitation due to emotional problems (3 items) and mental health (5 items) [16]. Items are scored on a varying 2- to 6-point Likert scale. The SF-36 also includes a single-item measure of health transition or change. The scores are summed per domain and then transformed to a 0–100 score (higher scores = better health). The first four domains can be summarized in the Physical Component Summary (PCS); the last can be summarized into the Mental Component Summary (MCS) [17]. A validated Dutch version of the SF-36 is available [18].

Statistical analysis

First, the percentages of missing values per item of the GIS were determined. Floor and ceiling effects were assessed for the total subscales of the GIS, the HAQ-DI and the eight different domains and summed scores of the SF-36 by calculating the percentage of respondents scoring minimum or maximum scores on each scale. Floor or ceiling effects were considered relevant when >15% of the respondents scored worst or best, respectively, on each scale. Internal consistency within each subscale of the GIS was tested using Cronbach's α and the Spearman–Brown prophecy formula to adjust for a 10-item scale. Cronbach's $\alpha$was acceptable when >0.70 (or >0.80 when adjusted to a 10-item scale) [19].

The content of the GAQ2.0 was evaluated by linking the GAQ2.0 to the ICF using the updated ICF linking rules of Cieza *et al.* [20] and comparing the content with the HAQ and SF-36 (that were linked using the same rules). Construct validity was assessed by correlating the scores of the subscales of the GIS with the eight categories of the HAQ-DI and its total score and with the eight domains and summed scores (PCS/MCS) of the SF-36 using Spearman's correlation. Correlations <0.29 were considered small, 0.30–0.49 were moderate and >0.50 were considered large [21].

Discriminative capacity was tested by determining the differences in GIS scores across clinical characteristics that reflect disease severity: (i) use of uric acid–lowering therapy (ULT); (ii) the presence or absence and amount of tophi [0 tophi ($n = 82$), 1–3 tophi ($n = 22$) and >3 tophi ($n = 22$)]; (iii) gout flares last year (yes/no); (iv) visual analogue scale (VAS) during a typical gout attack and (v) pain between gout attacks using analyses of variance. In order to improve the interpretation we repeated these analyses for the generic measures HAQ and SF-36 (PCS and MCS). We hypothesized that the above-mentioned characteristics influence GIS scores (e.g. worse GIS scores in patients with tophaceous gout) for disease-specific but not generic instruments. Test–retest reliability was assessed within a 4-week interval between measurements in a sample of 51 patients with stable disease using intraclass correlation coefficients (ICCs). ICCs >0.70 were considered acceptable [22]. Stable disease was defined as self-reported stable gout in the past 4 weeks. All statistical analyses were conducted using PASW Statistics 19.0 (IBM, Armonk, NY, USA).

Results

Clinimetric validation

A total of 126 patients [106 (84%) male, mean age 66.6 years (S.D. 10.4), range 42–89 years, mean disease duration 11.2 years (S.D. 10.6)] participated (53% of those were invited), of which 60 (48%) patients had tophaceous gout. Demographic and clinical data are reported in Table 1.

Items were missing in 19 (15%) GAQ2.0 questionnaires. Missing items were random, although we postulated that the gout impact (GI) subscale wellbeing during attack had the most missing items, due to the highest number of items to be scored in the subscale.

Table 1 Demographic and clinical data of the 126 patients included in the validation study

Characteristics	
Age, mean (S.D.) [range], years	66.6 (10.4) [42–89]
Male sex, *n* (%)	106 (84.1)
Education, *n* (%)	
Low (high school or less)	89 (70.6)
High (college or more)	37 (29.4)
Disease duration, mean (S.D.) [range], years	11.2 (10.6) [0.5–52]
Tophaceous gout ever, yes, *n* (%)	60 (47.6)
Number of tophi at examination, *n* (%) (*n* = 124)	
0	80 (64.5)
1–3	21 (17.0)
4–9	14 (11.2)
10+	9 (7.3)
Flares last 12 months, *n* (%) (*n* = 116)	
0	41 (32.5)
1–2	46 (36.5)
3–5	23 (18.2)
6–10	12 (9.6)
>10	4 (3.2)
Uric acid lowering therapy, yes, *n* (%)	86 (68.3)
HAQ-DI (0–3), mean (S.D.)	0.6 (0.6)
SF-36, mean (S.D.)	
PCS	39.1 (12.0)
MCS	49.9 (12.3)
Pain	
Typical gout attack, VAS, mean (S.D.)	8.6 (1.1)
Pain due to gout, between attacks, *n* (%)	
None of the time	42 (33.3)
A little of the time	18 (14.3)
Some of the time	31 (24.6)
Most of the time	20 (15.9)
All of the time	15 (11.9)
GISs, mean (S.D.) [range]	
Gout concern overall	53.8 (22.4) [0–100]
Gout medication side effects	45.2 (21.3) [0–100]
Unmet gout treatment needs	48.1 (13.8) [0–100]
Well-being during attack	45.0 (11.3) [16–75]
Gout concern during attack	44.7 (22.1) [0–100]

HAQ-DI: HAQ – disability index; SF-36: Short Form (36) Health Survey; PCS: Physical Component Summary; MCS: Mental Component Summary.

Floor (extreme impact/need) and ceiling effects (no problem/need) were negligible for the GIS (floor effects ranged from 0.9 to 5.3% and ceiling effects ranged from 0.0 to 2.6%). Floor effects (high disability) were not observed for the HAQ-DI, while ceiling effects (no disability) were observed in 23.1%. Floor effects for the SF-36 were observed for role limitation due to emotional and physical problems in 18.4 and 37.2% of the patients, respectively, while ceiling effects were also observed in role limitation due to emotional (64%) and physical (35.4%) role problems, but also in the domain social functioning (25.6%).

Internal consistency tested by Cronbach's α was sufficient to excellent ($\alpha = 0.83$–0.94) for four of the five GI scales (gout concern overall, unmet gout treatment need, well-being during attack and gout concern during attack) when adjusted to a 10-item scale. The internal consistency of the GI scale gout medication side effects was poor ($\alpha = 0.51$). Data are shown in Table 2.

Test–retest reliability analysis was assessed in a group of 51 of 55 (93%) patients who completed the questionnaire twice within a 4-week interval and who reported their gout had been stable in the past 4 weeks [44 male (80%), mean age 67 years (S.D. 10.0), mean disease duration 11.0 years (S.D. 10.4), 40% tophaceous gout]. The ICCs were sufficient (ICC = 0.73–0.86) for all GISs, except for unmet gout treatment need (ICC = 0.56) (Table 3), which was only moderate. Further analysis showed 41 of 51 patients scored equal or better in the retest questionnaire for unmet gout treatment need.

Content validity, assessed by linking the constructs within the items of the GAQ2.0 to the categories of the ICF classification (mapping) identified 34 constructs across the 31 items of the GAQ (see Supplementary Table S1, available at *Rheumatology* Online) that were linked to 12 ICF categories, while 4 constructs could either not be linked [health condition (number of gout attacks, other health conditions) or could not be specified (overall physical health, overall mental health, quality of life)]. Of these, four ICF categories addressed the component body functions, seven ICF categories the component activities and participation and one category environmental factors. A relatively high number of constructs ($n = 9$) referred to emotional functions (b152) that usually addressed the direct consequence of gout but could also relate to emotions concerning activities [I fear (b152) that I cannot continue my hobbies (d920)]. When comparing the content of the SF-36 and HAQ it is clear that the GAQ addresses no specific limitations in arm use or mobility, does not address vitality and has less emphasis on activities and more emphasis on emotional functions.

Construct validity (Table 4) showed that gout concern overall and gout concern during attack showed moderately positive correlations with the HAQ-DI ($r = 0.37$ and 0.32, respectively) and moderately negative associations with the PCS and MCS, respectively ($r = -0.37$). In addition, the GI subscale gout medication side effects correlated moderately with the MCS ($r = -0.34$). For all other subscales, correlations were low.

The discriminative validity (Table 5) of the GI subscales could not support the hypothesis that patients with tophaceous gout, with more tophi, with more pain during an attack or on ULT scored worse on any GIS. Importantly, generic measures (SF-36 and HAQ-DI) were unable to discriminate across these measures. On the other hand, and as expected, patients with gout flares last year and patients with more pain between gout attacks had significantly more gout concerns overall and tended to have more unmet gout treatment needs and reported worse physical HRQOL as measured by the SF-36.

Discussion

The present study shows that the Dutch version of the GAQ2.0 has aspects of validity that make it worthwhile

Table 2 Internal consistency analysis for Gout Impact scales

Gout Impact Scale (*n*)	Items (*n*)	Cronbach's α	Adjusted α to 10-item scale[a]
Gout concern overall (116)	4	0.86	0.94
Gout medication side effects (114)	2	0.17	0.51
Unmet gout treatment need (111)	3	0.60	0.83
Wellbeing during attack (107)	11	0.90	[b]
Gout concern during attack (116)	4	0.72	0.86

[a]Adjusted to 10-item scale using the Spearman–Brown prophecy formula. [b]Not calculated.

Table 3 Test–retest reliability ($n = 51$) of the subscales of the Gout Assessment Questionnaire

	ICC	95% CI	Significance (*P*)
Gout concern overall	0.82	0.70, 0.90	<0.001
Gout medication side effects	0.81	0.67, 0.89	<0.001
Unmet gout treatment need	0.56	0.25, 0.74	<0.001
Wellbeing during attack	0.86	0.74, 0.92	<0.001
Gout concern during attack	0.73	0.53, 0.84	<0.001

ICC: intraclass correlation coefficient.

Table 4 Construct validity Pearson correlations for Gout Impact scales with HAQ scores, HAQ-DI and SF-36

	Gout concern overall	Gout medication side effects	Unmet gout treatment needs	Well-being during attack	Gout concern during attack
Dressing, *r* (*n*)	0.39** (111)	0.29** (109)	0.15 (108)	0.25* (102)	0.28** (111)
Rising, *r* (*n*)	0.37** (111)	0.21* (109)	0.19* (108)	0.10 (102)	0.30** (111)
Eating, *r* (*n*)	0.19* (111)	0.16 (109)	0.21* (108)	0.16 (102)	0.22* (111)
Walking, *r* (*n*)	0.32** (111)	0.24* (109)	0.06 (108)	0.14 (102)	0.27** (111)
Hygiene, *r* (*n*)	0.24* (112)	0.10 (110)	0.07 (109)	0.13 (103)	0.21* (112)
Reaching, *r* (*n*)	0.26** (112)	0.23* (110)	0.06 (109)	0.13 (103)	0.21* (112)
Gripping, *r* (*n*)	0.16 (112)	0.08 (110)	0.13 (109)	0.12 (103)	0.21* (112)
Activity, *r* (*n*)	0.30** (112)	0.15 (110)	0.13 (109)	0.09 (103)	0.19* (112)
HAQ-DI, *r* (*n*)	0.37** (111)	0.25** (109)	0.16 (108)	0.19 (102)	0.32** (111)
SF-36					
Physical function, *r* (*n*)	−0.34** (110)	−0.27** (108)	−0.13 (107)	−0.22* (101)	−0.29** (110)
Role physical, *r* (*n*)	−0.28** (107)	−0.14 (105)	−0.01 (104)	−0.01 (98)	−0.28** (107)
Bodily pain, *r* (*n*)	−0.53** (111)	−0.31** (109)	−0.29** (108)	−0.02 (102)	−0.42** (111)
General health, *r* (*n*)	−0.43** (109)	−0.37** (107)	−0.23* (106)	−0.03 (100)	−0.27 (109)
Vitality, *r* (*n*)	−0.44** (109)	−0.32** (107)	−0.19 (106)	−0.04 (100)	−0.39** (109)
Social function, *r* (*n*)	−0.40** (111)	−0.32** (109)	−0.24* (108)	−0.11 (102)	−0.32** (111)
Role emotion, *r* (*n*)	−0.19* (108)	−0.27** (106)	0.00 (105)	0.08 (99)	−0.30** (108)
Mental health, *r* (*n*)	−0.30** (109)	−0.31** (107)	−0.23* (106)	−0.12 (100)	−0.42** (109)
Physical summary, *r* (*n*)	−0.37** (104)	−0.21* (102)	−0.21* (101)	−0.09 (95)	−0.26** (104)
Mental summary, *r* (*n*)	−0.23* (104)	−0.34** (102)	−0.10 (101)	0.01 (95)	−0.37** (104)

*$P < 0.05$; **$P < 0.01$. HAQ-DI: HAQ – disability index; SF-36: Short Form (36) Health Survey.

for further consideration as a disease-specific instrument to assess HRQOL in patients with gout.

Translation of the English version was performed following standard and internationally validated procedures and no cultural adaptations were needed [11, 12]. In further clinimetric testing an important advantage over the commonly used generic measures in gout, namely the lower frequency of either ceiling or floor effects when compared with the HAQ or SF-36, was found. This is important because floor and ceiling effects tend to reduce

Table 5 Discrimination of gout impact scales across clinical measures

	Using uric acid lowering therapy (yes/no)		Number of tophi: 0, 1–3, >3		Flares last year (yes/no)		Pain, typical gout attack (VAS)		Pain gout between attacks (VAS)	
	F	*P*	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*
GISs										
Gout concern overall	1.760	0.177	0.453	0.637	10.873	0.001**	1.061	0.386	3.692	0.007**
Gout medication side effects	0.086	0.917	0.549	0.579	3.422	0.067*	1.157	0.335	2.014	0.097*
Unmet gout treatment need	0.677	0.510	0.021	0.980	3.400	0.068*	0.156	0.978	2.446	0.051*
Well-being during attack	0.562	0.572	0.516	0.598	0.083	0.774	0.817	0.540	0.156	0.960
Gout concern during attack	2.023	0.137	0.462	0/630	1.123	0.292	0.790	0.559	1.659	0.165
Generic measures										
SF-36 PCS	2.241	0.142	0.741	0.479	0.436	0.510	1.165	0.332	2.395	0.055*
SF-36 MCS	0.002	0.961	0.222	0.802	0.183	0.670	0.420	0.834	0.168	0.954
HAQ-DI	0.291	0.591	1.376	0.257	0.071	0.790	0.695	0.628	1.768	0.147

*$P < 0.10$; **$P < 0.05$.GIS: Gout Impact Scale; HAQ-DI: HAQ – disability index; SF-36: Short Form (36) Health Survey; PCS: Physical Component Summary; MCS: Mental Component Summary.

responsiveness. The higher ceiling effects in the HAQ and SF-36 as found in gout are similar to those reported in the literature [14, 23, 24].

The internal consistency of the different subscales was considered sufficient for all GISs after adjusting for the number of items, except for the two-question subscale (probably explaining the lower Cronbach's α) gout medication side effects. Compared with the internal consistency in the validation study of the English GAQ2.0, Cronbach's α were slightly lower, but still acceptable (>0.80 in the adjusted analyses).

Four-week test–retest reliability was sufficient in all gout impact subscales except for the unmet gout treatment need, which was only moderate, in contrast to the original validation study. Construct validity of all GI subscales showed overall low correlations with the HAQ. Only the gout concern (overall and during attack) GI subscale had moderate correlations with the HAQ-DI (several subscales) and the mental and physical component scores of the SF-36.

As such, the overall disappointing correlations are not surprising, since several subscales of the GIS address concerns with disease while the SF-36 and HAQ-DI address impairments/limitations. In other words, the GIS reflects other aspects of health.

This interpretation is further supported by content comparison using the ICF, which revealed the GIS has a stronger emphasis on emotional functions and less on limitations in functioning (fine hand function, mobility) or physical activities compared with the HAQ and SF-36. It was interesting to notice that the GI subscales tended to discriminate only between patients with and without flares and to a lesser extent with different levels of pain between attacks. Chronic pain and the number of flares likely reflect a worse impact, as was also found by Hirsch *et al.* [10]. In this regard it was surprising the GI subscales did not discriminate between patients with or without tophi, nor the number of tophi. In this respect, the GI subscales were not better or worse than generic measures of HRQOL such as the HAQ and SF-36. Apparently neither the simple presence nor the number of tophi affects patients' HRQOL in our study. This is in line with a recent systematic review by Chandratre *et al.* [25] on HRQOL in gout, who concluded the relationship between tophi and HRQOL was not robust, reporting variable effects (worsening *vs* no effects) of tophi on HRQOL.

Limitations of this study were a relatively small sample size recruited in a university rheumatologic clinic, which has resulted in a sample with a high number of patients with tophaceous gout and on ULT, limiting its generalizability to patients with less severe disease. The influence of selection bias on GAQ2.0 validity is difficult to predict. As our study covers the full spectrum of disease, the slight overpresentation of worse disease will probably have no important influence on the clinimetric properties.

The sample size likely accounts for the slightly lower Cronbach's α and ICCs as compared with the original validation study. Although we used the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) criteria, we were unable to assess all aspects that are relevant in developing and validating a PRO. In particular, we were unable to assess minimal important change (MIC), minimal important difference (MID) or patient acceptable symptom state. Within the COSMIN criteria, however, no consensus on standards for assessing MIC or MID could be reached [26]. Furthermore, there is large variation in the interpretation of these aspects in the literature [27]. MIDs of the GIS were assessed earlier by Khanna *et al.* [28], showing differences were significantly important when differing between 5 and 8 points per scale.

To the best of our knowledge, the GAQ2.0 is the only disease-specific PRO assessing the impact of gout on functioning and health. PROs are becoming increasingly important in outcome research, because not every aspect of disease can be measured using biomedical findings. This is clearly applicable to gout, with its heterogeneous manifestations (from asymptomatic hyperuricaemia to

chronic tophaceous gout). This highlights the need for a disease-specific instrument to understand the impact of gout. In clinical care the GAQ2.0 (and more specifically the GIS) could serve as a screening tool to identify to what extent (and on which subscale) patients experience the impact of gout. Also, as physicians and patients have different views on disease severity and needs, application of the GIS in clinical practice might help to adjust the choice of pharmacological and non-pharmacological interventions to the patients' needs, which in turn could also improve adherence as well as overall health outcome and well-being [29]. Ultimately, it may have an effect on health care resource utilization and costs. Having said this, the issue of interpretability of the GAQ2.0, as for any PRO, in clinical practice and its final impact on health and resource utilization needs further exploration [30]. In conclusion, the Dutch GAQ2.0 is a gout-specific PRO instrument that measures HRQOL. The GAQ2.0 seems to measure different aspects of health than generic instruments, such as the HAQ-DI and SF-36. It is therefore promising to further explore its predictive value with regard to long-term health outcomes. Our study contributes to the further development and testing of the GAQ2.0 as a promising gap-filling instrument to measure HRQOL.

Rheumatology key messages

- A Dutch translation of the Gout Assessment Questionnaire 2.0 is now available. No cultural adaptations were needed.
- The Gout Assessment Questionnaire 2.0 subscales show good internal consistency, test–retest reliability and correlate better with gout-specific outcomes.

Acknowledgements

We thank Dr Jan Hirsch for her help in the translation process and for providing information on analysis of the Gout Assessment Questionnaire 2.0.

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

Supplementary data

Supplementary data are available at *Rheumatology* Online.

References

1 Richette P, Bardin T. Gout. Lancet 2010;375:318–28.

2 McQueen FM, Chana A, Dalbeth N. Mechanisms of joint damage in gout: evidence from cellular and imaging studies. Nat Rev Rheumatol 2012;8:173–81.

3 Khanna PP, Nuki G, Bardin T *et al*. Tophi and frequent gout flares are associated with impairments to quality of life, productivity, and increased healthcare resource use. Results from a cross-sectional survey. Health Qual Life Outcomes 2012;10:117.

4 Singh JA. Quality of life and quality of care for patients with gout. Curr Rheumatol Rep 2009;11:154–60.

5 Brook RA, Kleinman NL, Patel PA *et al*. The economic burden of gout on an employed population. Curr Med Res Opin 2006;22:1381–9.

6 Taylor WJ, Schumacher HR Jr, Singh JA *et al*. Assessment of outcome in clinical trials of gout. Rheumatology 2007; 46:1751–6.

7 Colwell HH, Hunt BJ, Pasta DJ *et al*. Gout Assessment Questionnaire: initial results of reliability, validity and responsiveness. Int J Clin Pract 2006;60:1210–7.

8 Boers M, Brooks P, Strand CV *et al*. The OMERACT filter for outcome measures in rheumatology. J Rheumatol 1998;25:198–9.

9 Hirsch JD, Lee SJ, Terkeltaub R *et al*. Evaluation of an instrument assessing influence of gout on health-related quality of life. J Rheumatol 2008;35:2406–2414.

10 Hirsch JD, Terkeltaub R, Khanna D *et al*. Gout-disease specific quality of life and the association with gout characteristics. Patient Relat Outcome Meas 2010;1:1–8.

11 Wild D, Grove A, Martin M *et al*. Principles of good practice for the translation and cultural adaptation process for patient reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. Value Health 8:94–104.

12 Beaton DE, Bombardier C, Guillemin F *et al*. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 2000;25:3186–91.

13 Fries JF, Spitz P, Kraines RG *et al*. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137–45.

14 Alvarez-Hernández E, Peláez-Ballestas I, Vázguez-Mellado J *et al*. Validation of the health assessment questionnaire disability index in patients with gout. Arthritis Rheum 2008;59:665–9.

15 Bijlsma JW, Oude Heuvel CH, Zaalberg A. Development and validation of the Dutch questionnaire capacities of daily life (VDF) for patients with rheumatoid arthritis. J Rehab Sci 1990;3:71–4.

16 Ware JE, Sherbourne CD. The RAND-36 Short-Form Health Status Survey: 1. Conceptual framework and item selection. Med Care 1992;30:473–81.

17 Ware JE, Kosinski M, Dewey JE. How to Score Version 2 of the SF-36 Health Survey. Lincoln, RI, USA: QualityMetric, 2000.

18 Aaronson NK, Muller M, Cohen PD *et al*. Translation, validation and norming of the Dutch language version of the SF-36 health survey in community and chronic disease populations. J Clin Epidemiol 1998;51:1055–68.

19 Nunnally JC. Psychometric Theory, 2nd edn. New York: McGraw-Hill, 1978.

20 Cieza A, Geyh S, Chatterii S *et al*. ICF linking rules: an update based on lessons learned. J Rehabil Med 2005;37: 212–8.

21 Cohen J, Cohen P, West SG *et al*. Applied Multiple Regression/Correlation Analysis for the Behavioural Sciences. 2nd edn. Hillsdale, NJ: Lawrence Erlbaum, 2003.

22 Fitzpatrick RD, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess 1998;2:1–74.

23 Taylor WJ, Colvine K, Gregory K *et al*. The health assessment questionnaire disability index is a valid measure of physical function in gout. Clin Exp Rheumatol 2008;26: 6206.

24 Ten Klooster PM, Oude Voshaar MA, Taal E *et al*. Comparison of measures of functional disability in patients with gout. Rheumatology 2011;50:709–13.

25 Chandratre P, Roddy E, Clarson L *et al*. Health-related quality of life in gout: a systematic review. Rheumatology 2013;52:2031–40.

26 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.

27 Terwee CB, Roorda LD, Dekker J *et al*. Mind the MIC: large variation among populations and methods. J Clin Epidemiol 2010;63:524–34.

28 Khanna D, Sarkin A, Khanna P *et al*. Minimally important differences of the gout impact scale in a randomized controlled trial. Rheumatology 2011;50:1331–6.

29 Sarkin AJ, Levack AE, Shieh MM *et al*. Predictors of doctor-rated and patient-rated gout severity: gout impact scales improve assessment. J Eval Clin Pract 2010;16: 1244–7.

30 Higginson IJ, Carr AJ. How can quality of life measures be used in the clinical setting? BMJ 2001;322:1297–300.