doi: 10.1111/j.1742-1241.2006.01104.x

Gout Assessment Questionnaire: initial results of reliability, validity and responsiveness

H. H. COLWELL,[1] B. J. HUNT,[2] D. J. PASTA,[1] W. A. PALO,[2] S. D. MATHIAS,[1] N. JOSEPH-RIDGE[2]
[1]*Ovation Research Group, San Francisco, CA,* [2]*TAP Pharmaceutical Products Inc., Lake Forest, IL, USA*

SUMMARY

We evaluated the psychometric properties of a new gout-specific patient reported outcomes questionnaire. The Gout Assessment Questionnaire (GAQ) and the SF-36 were administered to 126 subjects in a multicenter Phase II program of febuxostat, an investigational treatment for hyperuricemia (serum uric acid >8.0 mg/dl) in patients with chronic gout. The questionnaire was administered at baseline and 1, 6 and 12 months later. The majority of subjects, mean age 54 years, were male, Caucasian and had experienced a gout flare within the last year. Seven domains were identified, all met criteria for reliability and validity. Cronbach's alpha ranged from 0.78 to 0.97. Pearson correlations between GAQ and SF-36 scales were generally low to moderate, with the highest correlation between Gout Pain and Severity and SF-36 Bodily Pain, $r = 0.45$. Guyatt's statistic (measure of responsiveness) ranged from 0.24 to 1.00 at 12 months. Minimal clinically important differences ranged from 2 (Gout Concern) to 10 points (Productivity). The GAQ has acceptable psychometric properties. Further research is required to confirm results, which may provide more information to improve the GAQ for use in clinical trials.

Keywords: Gout; patient reported outcomes; questionnaire development; psychometrics; SF-36; Gout Assessment Questionnaire

INTRODUCTION

Gout is a disease resulting from the deposition of monosodium urate crystals in the joints and soft tissues (1). Chronic gout is accompanied by acute gout flares, which are episodic, intense inflammatory responses to the deposits of urate crystals in the joints and other soft tissues. Symptoms include sudden onset of severe pain, stiffness, inflammation and limited range of motion in the affected joints (2). While most initial gout attacks are monoarticular with the inflammation of the big toe, other involved joints may include the ankle, wrist and fingers (3). Initial symptoms of a gout flare are usually self-limiting, but in some patients, the flare symptoms may take up to a week or more to subside. Left untreated for many years, urate crystal deposition can result in a chronic deforming arthritis; the development of crystal aggregates (tophi), which can cause destruction of cartilage and bone; and, on occasion, organ dysfunction, especially renal impairment because of uric acid nephrolithiasis (4).

Correspondence to:
Barbara J. Hunt, TAP Pharmaceutical Products Inc, 675 North Field Drive, Lake Forest, IL 60045
Tel.: + 847 582 5677
Fax: + 847 582 5600
Email: barbara.hunt@tap.com

Prevalence of Gout

In the 1996 US National Health Interview Survey, 2.24% of people between 45 and 64 years of age reported gout, while 3.08% of people 65 years and older reported gout (5). While hyperuricemia is a major risk factor, not all patients with hyperuricemia have gout and gout can occur in those with normal urate levels (6). Hypertension has been associated with gout, as has obesity, lead exposure, diuretic use, alcohol consumption, renal insufficiency and family history (7). Between 1977 and 1995, the incidence of gout has risen twofold from 35.1/100,000 to 56.4/100,000. The overall male to female ratio remained the same and there was no difference in the prevalence of comorbid conditions of diabetes, hypertension, coronary artery disease or hypothyroidism. While the increase may be a result of improved ascertainment, it may also be related to other risk factors (8).

The impact of gout from the patient's perspective appears to be largely unexplored. A Medline literature search conducted in 2000 and again in 2006 did not reveal the existence of any patient-reported outcomes (PRO) questionnaires for use in individuals with gout. This is in contrast to other rheumatological conditions such as osteoarthritis and rheumatoid arthritis where multiple PRO measures are available (9). The value of PRO questionnaires is no longer debatable, as such measures have become an

integral component of clinical trials. Gout is no exception to these other conditions. There is a huge potential value for a validated measure to provide insight from the patient's perspective. The OMERACT initiative (Outcome Measures in Rheumatology) has been a major proponent of the importance of considering the patient perspective in the treatment of rheumatological conditions (10). The OMERACT gout special interest group has proposed a list of domains to be included in chronic and acute gout trials, including health related quality of life and functioning; however, these proposed domains require additional evaluation prior to gaining full endorsement by OMERACT (11). In June 2004, the FDA arthritis advisory board held a meeting to discuss types of outcomes to be captured in gout clinical trials; patient reported outcomes were discussed as an important component (12).

Febuxostat is a novel non-purine selective inhibitor of xanthine oxidase under development for the treatment of hyperuricemia (serum uric acid >8.0 mg/dl) in patients with chronic gout (13–15). Therefore, it was important to develop a tool to assess the quality of life in gout patients.

Objective

The objective of this study was to assess the reliability, validity and responsiveness of a newly developed gout-specific PRO questionnaire, which could be used in clinical trials and clinical practice to more fully explore the impact of gout and its treatment from the patient's perspective.

SUBJECTS AND METHODS

A detailed MedLine literature search, conducted in 2000 for the years 1966–2000 inclusive, failed to identify any gout-specific patient reported outcomes instruments. However, the review was very helpful in identifying the most common symptoms experienced by gout patients and how they might impact a patient. Based on these findings, a draft questionnaire was developed, which included items assessing pain, gout flares, health distress, days lost from work and treatment satisfaction. Three rheumatologists (all with academic appointments; one also in private practice and another also at a Veterans Administration Hospital) who treat patients with gout and an expert in patient education and assessment, were interviewed and asked to provide feedback via telephone on the draft questionnaire. Minor modifications were made to the questionnaire based on their feedback. Interviews were also conducted with five adult gout patients. Each patient received a copy of the questionnaire by mail and was asked to complete it prior to taking part in an in-depth telephone interview. Based on the feedback from these five patients, the questionnaire was further revised and finalised. The final Gout Assessment Questionnaire (GAQ) contained 21 items assessing a variety of gout-specific aspects of health including pain, well-being, productivity and treatment satisfaction. Responses were generally Likert-type scales (e.g. none of the time to all of the time; or not at all to extremely), with an exception being the activity restriction items which asked for specific numbers of days or hours of limitation. Questionnaire items are summarised in Table 1.

The GAQ, along with the SF-36 (v 1) (16) and a modified version of the Medical Outcomes Study (MOS) Health Distress scale (17), were administered as part of the Phase II program evaluating the safety and efficacy of febuxostat (Figure 1). In the initial Phase II trial, subjects were asked to complete the questionnaires at baseline (Day 1) and at the end of the 1-month study period (or upon withdrawal from the study). (Questions about treatment bother, satisfaction and convenience were not included as part of the baseline questionnaire, as they were not relevant.) Subjects who completed the study were offered the option of enrol-

Table 1 Scales and Number of Items in the Gout Assessment Questionnaire

Scale	*No. of items*	*Question example*
Gout Concern	6	I am worried that I will have a gout attack or flare within the next year.
Well-Being	6	During the most recent flare, how much did your symptoms interfere with your mood?
Productivity	2	Think about your most recent attack of gout or what is also called a gout flare. How many hours or days did you miss work or were unable to complete your household or family responsibilities because of this gout flare?
Gout Pain and Severity	2	Some people experience gout pain or discomfort regularly even when they are NOT experiencing a gout flare. In the time since you last completed this questionnaire, how much of the time did you experience pain related to your gout? Do not include pain relating to a gout flare.
Treatment Convenience	1	How convenient is your study medication?
Treatment Satisfaction	3	How likely is it that you would continue to use this study medication for your gout?
Treatment Bother	1	How bothered were you by side effects from your study medication for your gout (e.g. diarrhoea, rash, headache, etc.)?

All scales are scored on a 0–100 scale with a higher score indicating better functioning or higher satisfaction.

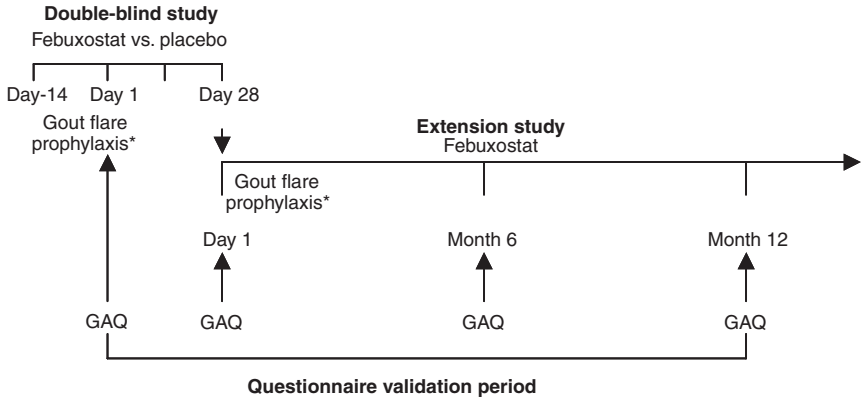

Figure 1 Febuxostat Phase 2 study diagram and administration of Gout Questionnaire. GAQ = Gout Assessment Questionnaire administration. Note: Questionnaire was administered at scheduled visit or at the time of discontinuation. *Gout flare prophylaxis was provided from Day-14 to Day 1 if washing out of previous urate-lowering therapy. All subjects received prophylaxis for the first 2 weeks of the double-blind study and for the first 4 weeks of the extension study. Gout flares were treated as needed, regimen per investigator

ling in an open-label extension study, in which the questionnaire was completed at 6 and 12 months. These data were used to develop the scale structures, as well as to assess reliability, validity and responsiveness of each of the scales.

Study Populations

The main study population for the scale creation and questionnaire validation analyses included all enrolled subjects from the 4-week study (enrollment criteria included a serum uric acid (sUA) >8.0 mg/dl) who had completed at least one item from the Day 1 questionnaire. This cohort was termed the full cohort. One set of validity analyses was performed on a subset of the full cohort called the responsiveness cohort. Several responsiveness cohorts were generated, depending on the time point to be analysed. One responsiveness cohort was comprised of subjects in the full cohort who had a non-missing baseline and a non-missing sUA level at 1 month in the 4-week study. Additionally, separate cohorts were developed for subjects in the full cohort who had non-missing baseline values and non-missing sUA levels at 6 and 12 months in the extension study.

Developing Scales

The creation of scales for the GAQ was accomplished by performing a series of analyses, including variable clustering as implemented in the SAS VARCLUS procedure (18). The approach used iterative splitting and factor analytic methods to divide the group of variables from the GAQ into discrete (non-overlapping) subgroups that were relatively highly correlated and well-represented by a single scale value. The items included in the variable clustering were first rescaled from 0 to 100 based on the possible responses and reversed if necessary so that a higher score was indicative of better health.

Analyses were conducted using all of the data from all available study subjects (full cohort). All rescaled GAQ items were included in the first variable clustering procedure. In subsequent variable clustering procedures, groups of items expected to be related were clustered separately. Variables that did not fit well with any scale from an empirical standpoint were either assigned on theoretical grounds or omitted from all clusters. The weighting of the variables was examined after establishing which items belonged within each scale. As the proportion of the variance explained by the simpler, unweighted mean of the items was nearly as great as for the optimally-weighted mean, we used the unweighted mean of the items to form the scale. When scales were developed, a value for the scale was calculated if at least one half of the items within the scale were available. Seven scales were identified: Gout Concern, Well-Being, Productivity, Gout Pain and Severity, Treatment Satisfaction, Treatment Convenience and Treatment Bother.

Reliability and Validity Analyses

Reliability and validity were assessed in a number of ways. Internal-consistency reliability, a measure of the extent to which items within each scale correlate with each other to form a multi-item scale, was assessed using Cronbach's alpha. Data from all assessments were used to evaluate internal-consistency reliability. An alpha coefficient of 0.70 or greater was considered acceptable (19). Construct validity was evaluated by calculating Pearson correlation coefficients among items and scales and determining the extent to which hypothesised relationships were upheld. We hypothesised that:

1 Subjects who report more severe pain would report worse physical functioning (SF-36) and worse Bodily Pain (SF-36) and less treatment satisfaction than those who reported less severe pain.

2 Subjects who reported more frequent pain would report worse physical functioning (SF-36) and worse Bodily Pain (SF-36) and less treatment satisfaction than those who reported less frequent pain.
3 Gout Concern should be at least moderately correlated with health distress.
4 Subjects who meet the efficacy definition (in terms of achieving a sUA level of <6.0 mg/dl) should report higher treatment satisfaction than those who did not reduce their sUA to efficacious levels.

Known-groups validity was assessed by categorising the study subjects according to their sUA levels at the time point to be analysed (i.e. at 1 month, 6 months and 12 months). Three groups were created: those whose sUA level improved to <6.0 mg/dl, those who improved to ≥ 6.0 to <7.8 mg/dl and those who had ≥ 7.8 mg/dl. Mean scale scores were compared for each GAQ and SF-36 scale and MOS Health Distress at 1, 6 and 12 months using analysis of variance (ANOVA) with the Tukey adjustment for multiple comparisons. It was hypothesised that the group with the lowest sUA levels would report better functioning on each of the GAQ scales compared with the group with the highest sUA levels.

Responsiveness and Minimally Clinically Important Difference

Responsiveness is designed to evaluate how effectively the questionnaire detects change in those individuals who are known to change clinically. Responsiveness is demonstrated if the change scores improve, on average, for the study subjects demonstrating clinical improvement. Guyatt's statistic is calculated as the ratio of the average change score for an item or scale of the clinically improved study subjects to the standard deviation of the change score for that same item or scale of the clinically stable study subjects. A Guyatt's statistic with a magnitude of 0.20 indicates acceptable responsiveness, while a statistic with a magnitude of 1.00 or greater is considered indicative of a highly responsive item or scale (20). In this case, improved and stable subjects were defined using sUA levels.

The minimal clinically important difference (MCID) for each of the scales was also investigated. MCID was calculated using the patient-reported pain frequency and pain severity items from the GAQ as anchors. These items were used to assess the amount of change in the other scales associated with a minimum 1-point change in the pain scale (of a 5-point scale), using linear regression. Change scores were evaluated from baseline to 1 month, from baseline to 6 months and from baseline to 12 months for all non-pain GAQ scales included in both the baseline and follow-up questionnaires. All statistical analyses were performed using SAS software, Version 8.2 of the SAS System for Windows (18).

RESULTS

A total of 153 study subjects at 24 sites enrolled in the initial month-long Phase II randomised controlled trial of febuxostat. Of the 153 subjects enrolled, 126 (82.4%) completed at least one GAQ and were included in these analyses. The demographics of the full cohort are provided in Table 2. The majority of study subjects were male, Caucasian and had experienced a gout flare within the last year. The mean age of the study subjects was 54 years. The majority were current alcohol drinkers, but were not current tobacco users.

Table 3 contains the inter-scale Pearson correlations and the Cronbach's alpha values for the GAQ scales. The lower triangle displays the inter-scale correlations for the GAQ, while the Cronbach's alpha values are shown in the last column. Internal consistency reliability was demonstrated for the scales, as the Cronbach's alpha values ranged from 0.78 to 0.97, all above the acceptable level of 0.70.

Table 4 contains the inter-scale Pearson correlations between the GAQ and the SF-36. The SF-36 domain of Bodily Pain tended to have the strongest inter-scale correla-

Table 2 Demographics

Characteristic	*Full cohort (n = 126)*
Gender (%)	
Male	111 (88.1)
Female	15 (11.9)
Race (%)	
Caucasian	109 (86.5)
Black	8 (6.3)
Asian	2 (1.6)
Hispanic	4 (3.2)
Other	3 (2.4)
Age (years)	
Mean (SD)	53.7 (12.92)
Range	23–80
Tobacco use (%)	
Non-tobacco user	53 (42.1)
Ex-tobacco user	48 (38.1)
Tobacco user	25 (19.8)
Alcohol use (%)	
Non-drinker	28 (22.2)
Ex-drinker	14 (11.1)
Drinker	84 (66.7)
Years since gout diagnosis (%)	
>10 years ago	54 (42.9)
5–10 years ago	29 (23.0)
1–5 years ago	31 (24.6)
<1 year ago	12 (9.5)
Years since last gout flare (%)	
5–10 years ago	3 (2.4)
1–5 years ago	17 (13.5)
<1 year ago	106 (84.1)

SD = standard deviation.

Table 3 Construct validity (Inter-scale Pearson correlations) and internal consistency reliability (Cronbach's alpha) within GAQ scales on the full cohort

	Pearson's correlation coefficient between scales							
GAQ Scale	*Well-Being (WB)*	*Productivity (P)*	*Gout Concern (GC)*	*Treatment Satisfaction (TS)*	*Gout Pain and Severity (PS)*	*Treatment Bother (TB)*	*Treatment Convenience (TC)*	*Cronbach's alpha**
WB								0.97
P	0.69							0.90
GC	0.46	0.40						0.90
TS	0.23	0.22	0.39					0.78
PS	0.40	0.38	0.45	0.28				0.83
TB	0.01	0.10	0.06	0.18	0.12			N/A
TC	−0.04	0.02	0.25	0.30	0.05	0.22		N/A

GAQ, Gout Assessment Questionnaire. *Cronbach's alpha was not calculated (N/A) for single-item scales.

Table 4 Construct validity (Inter-scale Pearson correlations) between SF-36, Health Distress and GAQ scales on the full cohort

SF-36 domains	*Well-Being*	*Productivity*	*Gout Concern*	*Treatment Satisfaction*	*Gout Pain and Severity*	*Treatment Bother*	*Treatment Convenience*
Physical functioning	0.17	0.27	0.18	0.16	0.24	0.17	−0.01
Role – physical	0.25	0.30	0.31	0.28	0.25	0.20	0.04
Bodily Pain	0.30	0.29	0.41	0.35	0.45	0.17	0.04
General health	0.17	0.18	0.31	0.08	0.18	0.11	0.07
Vitality	0.13	0.21	0.27	0.11	0.22	0.20	0.14
Social functioning	0.25	0.26	0.34	0.26	0.28	0.13	0.01
Role – emotional	0.13	0.16	0.13	0.02	0.14	0.18	0.11
Mental health	0.17	0.16	0.29	0.00	0.21	0.08	0.02
Health transition	0.13	0.16	0.20	0.30	0.19	0.14	0.18
Summary							
Physical component summary	0.27	0.33	0.34	0.31	0.34	0.20	0.02
Mental component summary	0.13	0.13	0.23	−0.01	0.15	0.12	0.09
Medical Outcomes Study							
Health distress	0.28	0.29	0.46	0.17	0.29	0.11	0.03

tions with the GAQ scales for Well-Being, Productivity, Gout Concern and Gout Pain and Severity, ranging from 0.29 for Productivity to 0.45 for Gout Pain and Severity. The Role-Physical and Social Functioning domains also had correlations with these GAQ scales in the 0.25–0.34 range. As hypothesised, Gout Concern was highly correlated with MOS Health Distress (0.46). In general, the GAQ scales were more highly correlated with one another than with the scales of the SF-36.

It was hypothesised that subjects reporting more severe and more frequent pain would report worse physical functioning, worse Bodily Pain and less satisfaction than those with less severe pain. This was confirmed as moderate correlations were reported with both pain severity (0.23 for Physical Functioning, 0.44 for Bodily Pain and 0.28 for Treatment Satisfaction) and pain frequency (0.21 for Physical Functioning, 0.38 for Bodily Pain and 0.24 for Treatment Satisfaction) (data not shown).

Known-groups validity for the study population at 1 month compared three different levels of sUA groups. However, at 6 and 12 months no subjects had sUA levels $\geq$7.8 mg/dl, so only two groups were compared. No statistically significant differences between the groups emerged for the three sUA groups at 1 month, nor for the two groups at 6 and 12 months. Not surprisingly, the groups reported similar SF-36 scores at all time points (data not shown).

Table 5 contains the responsiveness results for the GAQ scales at 1, 6 and 12 months. With the exception of Gout Concern at 1 month (0.030), the Guyatt's statistics indicate responsiveness for all other scales for all time points considered with values ranging from 0.24 (Gout Pain and Severity at 12 months) to 1.14 (Well-Being at 6 months).

The mean change in GAQ scales was compared by change in pain severity and frequency. Results were similar for severity and frequency at all time points. Table 6 displays the mean

Table 5 Responsiveness of GAQ scales: mean GAQ change scores by serum uric acid level at follow up

	Serum uric acid level						
	Clinically improved (< 6.0 mg/dl)		*Clinically stable (≥6.0 to < 7.8 mg/dl)*		*Clinically worse (≥7.8 mg/dl)*		
GAQ Scale	*n*	*Mean change (SD)*	*n*	*Mean change (SD)*	*n*	*Mean change (SD)*	*Guyatt's Statistic**
Month 1							
Well-Being	61	32.46 (36.32)	21	32.82 (39.65)	30	35.14 (35.08)	0.819
Productivity	51	33.80 (43.07)	19	44.93 (42.03)	28	21.45 (51.32)	0.804
Gout Concern	63	0.60 (15.93)	21	5.95 (19.57)	29	6.24 (12.77)	0.030
Gout Pain and Severity	60	5.21 (25.33)	21	17.02 (21.07)	29	10.52 (25.93)	0.247
Month 6†							
Well-Being	60	34.69 (35.62)	8	31.77 (30.37)			1.142
Productivity	50	29.12 (46.27)	8	14.97 (66.16)			0.440
Gout Concern	60	15.97 (19.71)	8	14.38 (18.51)			0.863
Gout Pain and Severity	60	11.54 (28.53)	8	17.81 (36.70)			0.314
Month 12†							
Well-Being	49	31.82 (33.73)	12	29.17 (38.31)			0.831
Productivity	42	31.98 (47.39)	11	36.43 (46.64)			0.686
Gout Concern	49	20.61 (23.64)	12	25.69 (20.63)			0.999
Gout Pain and Severity	47	9.15 (22.92)	12	24.38 (38.51)			0.238

GAQ, Gout Assessment Questionnaire; SD, standard deviation. Higher mean change scores indicate better functioning. *Guyatt's statistic for each scale defined as the ratio of the mean change for the subjects who are clinically improved to the standard deviation of the change for subjects who are clinically stable. †No subjects had serum uric acid levels ≥7.8 mg/dl at 6 or 12 months.

Table 6 Mean change in non-pain-related GAQ scales based on the change in pain severity from baseline to 6 months

	*Change in pain severity from baseline to 6 months**						
	Improved		*Same*		*Worsened*		
GAQ Scale	*n*	*Mean change (SD)*	*n*	*Mean change (SD)*	*n*	*Mean change (SD)*	*Guyatt's Statistic†*
Well-Being	30	41.14 (41.15)	20	28.83 (29.64)	16	28.39 (25.65)	1.39
Productivity	28	42.53 (50.20)	16	18.24 (45.74)	12	−0.31 (37.90)	0.93
Gout Concern	30	18.53 (20.55)	20	12.21 (16.52)	16	16.67 (21.57)	1.12

GAQ, Gout Assessment Questionnaire; SD, standard deviation. Higher mean change scores indicate better functioning. *Improved: ≥1 category improvement from baseline; same: no change from baseline; worsened: ≥1 category worsening from baseline. †Guyatt's statistic for each scale defined as the ratio of the mean change for the subjects who are clinically improved to the standard deviation of the change for subjects who are clinically stable.

change score for the non-pain GAQ scales when subjects were categorised by the extent of improvement in pain severity from baseline to 6 months, as an example.

As expected, greater positive changes were noted for the Improved group (at least one category improvement from baseline) compared with the Worsened group (at least one category worsening from baseline). The Improved and Worsened group reported similar change scores in pain severity at 12 months for Well-Being and Gout Concern (data not shown).

Table 7 reports the MCID for each non-pain scale of the GAQ using pain frequency and severity as anchors (all time points included). With the exception of Well-Being for pain frequency, the MCID for each scale was statistically significantly >0. This implies Well-Being may not be substan-

Table 7 Minimal clinically important difference in GAQ scales linear regression on change in pain frequency and change in pain severity from baseline (all time points included)

GAQ Scale	*Beta (95% CI)**	*p-value†*
Pain frequency		
Well-Being	3.06 (−1.49, 7.60)	0.186
Productivity	12.33 (5.86, 18.81)	<0.001
Gout Concern	2.73 (0.11, 5.34)	0.041
Pain severity		
Well-Being	4.94 (2.34, 7.54)	<0.001
Productivity	8.25 (4.55, 11.96)	<0.001
Gout Concern	1.88 (0.32, 3.44)	0.018

GAQ, Gout Assessment Questionnaire. *Parameter estimate and 95% confidence interval for change in pain frequency/severity. †p-value for parameter estimate for change in pain frequency/severity.

tially affected by change in pain frequency. In interpreting the table, consider the Productivity scale. On average, a 10-point change is necessary on the scale for the change to be clinically meaningful.

DISCUSSION

The goal of this study was to evaluate the initial reliability, validity and responsiveness of the GAQ. In general, the results suggest the GAQ has acceptable properties of reliability and validity in this gout population. The questionnaire has excellent internal consistency, as demonstrated by the Cronbach's alpha values for all GAQ scales. In fact, most scales have alpha values over 0.90. The construct validity of the instrument was supported by the inter-scale correlations between the GAQ scales, the SF-36 and MOS Health Distress scale. As expected, correlations were higher between scales assessing similar concepts than between scales of dissimilar concepts. For example, the correlations between the GAQ scales tended to be moderate to high, while the correlations between the GAQ and the SF-36 tended to be low to moderate.

There are some areas that require further consideration. Our hypothesis that study subjects with a sUA level of <6.0 mg/dl would report higher treatment satisfaction than those who did not reduce their levels as much was not confirmed. As noted in Table 5, the majority of subjects were clinically improved (i.e. achieved sUA levels of <6.0 mg/dl) at the follow-up visits, limiting the ability of the questionnaire to distinguish between study subjects categorised cross-sectionally according to sUA levels in this study. Although no statistically significant differences were reported between the <6.0 mg/dl sUA group and the ≥ 6.0 to <7.8 mg/dl sUA group, some of the scales did show trends in the expected direction (i.e. higher functioning in the <6.0 mg/dl sUA group). Additional clinical measures such as physician assessment should be considered for further evaluation of known-groups validity. Initiation of therapy with a sUA lowering agent may result in more acute gout flares. Consequently, improvement in patient-reported outcomes may not be noted until after at least 6 months of treatment when the sUA has a chance to stabilise and total body urate load is decreasing. This is one reason to evaluate longer term data (6 months and 1 year) when considering the impact of gout treatments from the patient's perspective.

For the most part, the GAQ scales were responsive to change, the only scale that appeared unresponsive was Gout Concern at 1 month. This is not surprising considering the issues addressed in the Gout Concern scale, such as fear of gout flares, would probably take more than 1 month to be impacted by efficacious treatment. The majority of the study subjects had been diagnosed with gout many years ago and had experienced a flare within the past year. It is likely that they might require more time until they have full confidence that the treatment is working, manifesting in less Gout Concern for the future. Most of the improvement in Well-Being and Productivity scores occurred by 1 month, as noted by the relatively stable change scores at each subsequent time point after 1 month. The scales for Gout Concern and Gout Pain and Severity generally continued to improve at 6 and 12 months.

One limitation of the current questionnaire is that the MCIDs suggest that small changes will be clinically important with measures such as Gout Concern, while larger changes are needed with other measures such as Productivity. This should be addressed in the continued development of the questionnaire. The results from the MCID analyses suggest that a clinically meaningful change on the questionnaire could be as low as 2 points (Gout Concern) and as high as approximately 10 points (Productivity). In general, although all groups improved over time (including the clinically stable group), the improved group reported large improvements in most domains as early as 1 month. These results should be interpreted with caution given the small sample sizes and the fact that an objective, clinical measure was not available as the reference. Subjects in this study were fairly homogeneous, with the majority being male, Caucasian and having experienced a gout flare within the past year; therefore this study population was fairly consistent with the gout population as a whole. However, several of the analyses were conducted with small sample sizes and therefore more definitive results might be obtained with a larger sample.

Furthermore, the importance of symptoms to individual patients with gout varies widely. The ability of urate-lowering therapy to lower sUA levels below the level of physiologic saturation can cause a paradoxical increase in the number of flares experienced by patients when starting a new therapy or changing dosages of current therapy. The ability of a questionnaire to address these issues needs further evaluation. This questionnaire focused on the patient's perspective about their health; the ability of a questionnaire to distinguish between the patient's perspective and the physician's opinion of the patient's health is another area that needs further research.

In summary, initial results suggest the GAQ is reliable, valid and responsive in this gout population. While the results are encouraging, additional work on the questionnaire is required. Ongoing research includes further validation efforts to confirm scale structure, modification of items to eliminate open-ended questions and examining the questionnaire's reproducibility. Confirmatory analyses conducted on GAQ data in future studies will provide additional support to the psychometric properties of the questionnaire.

ACKNOWLEDGEMENTS

This project was funded by TAP Pharmaceutical Products Inc. The authors would like to acknowledge Connie Chen, PharmD and Katrine Wallace, PhD for their contributions to the development and validation of the questionnaire. The authors would also like to thank Jane T. Osterhaus, PhD for her knowledge, expertise and clinical perspective regarding the questionnaire and development of the manuscript; and Susan Cazzetta, RPh, Patricia MacDonald, NP and Pankaj Patel, PharmD for their editorial assistance in the preparation of the manuscript.

REFERENCES

1 Kim KY, Schumacher HR, Hunsche E et al. A literature review of the epidemiology and treatment of acute gout. *Clin Ther* 2003; **25**: 1593–617.
2 Perkins P, Jones AC. Gout. *Ann Rheum Dis* 1999; **58**: 611–7.
3 Harris MD, Siegal LB, Alloway JA. Gout and hyperuricemia. *Am Fam Physician* 1999; **59**: 925–34.
4 Argawal AK. Gout and pseudogout. *Prim Care* 1993; **20**: 839–55.
5 Adams PF, Hendershot GE, Marano MA. *Current estimates from the national health interview survey. National Center for Health Statistics*. Vital Health 18 Stat 1999 [cited 2004 Feb 11];10(200):1–212. Available from: URL: http://www.cdc.gov/nchs/data/series/sr_10/sr10_200.pdf., 1996.
6 Campion EW, Glynn RJ, DeLabry LO. Asymptomatic hyperuricemia: risks and consequences in the Normative Aging Study. *Am J Med* 1987; **82**: 421–6.
7 Roubenoff R, Klag MJ, Mead LA et al. Incidence and risk factors for gout in white men. *JAMA* 1991; **266**: 3004–7.
8 Arromdee E, Michet CJ, Crowson CS et al. Epidemiology of gout: is the incidence rising? *J Rheumatol* 2002; **29**: 2403–6.
9 Wolfe F, Lassere M, van der Heijde D et al. Preliminary core set of domains and reporting requirements for longitudinal observational studies in rheumatology. *J Rheumatol* 1999; **26**: 484–9.
10 Kvien T, Heiberg T. Patient perspective in outcome assessments – perceptions or something more? *J Rheumatol* 2003; **30**: 873–6.
11 Schumacher HR Jr, Edwards LN, Perez-Ruiz F et al. Outcome measures for acute and chronic gout. *J Rheumatol* 2005; **32**: 2452–5.
12 Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research arthritis advisory committee. Transcript. Wednesday, June 2, 2004 via http://www.fda.gov/ohrms/dockets/ac/04/transcripts/2004–4044T1.htm (accessed on January 31 2006).
13 Becker MA, Schumacher HR Jr, Wortmann RL et al. Febuxostat, a novel nonpurine selective inhibitor of xanthine oxidase: a twenty-eight-day, multicenter, phase II, randomized, double-blind, placebo-controlled, dose-response clinical trial examining safety and efficacy in patients with gout. *Arthritis Rheum* 2005; **52**: 916–23.
14 Schumacher HR, Wortmann R, Becker MA et al. A phase 2, long term open-label safety and efficacy study of febuxostat, a novel non-purine, selective inhibitor of xanthine oxidase [abstract]. *Arthritis Rheum* 2004; **50** (9 Suppl.): S335.
15 Schumacher HR, Wortmann R, Becker M et al. Long-term safety and efficacy of febuxostat, a novel non-purine selective inhibitor of xanthine oxidase, in subjects with hyperuricemia and gout [abstract]. *Ann Rheum Dis* 2005; **64** (Suppl. III): 498. Abstract no. SAT0282.
16 Ware JE, Snow KK, Kosinski M, Gandek B. *SF-35 Health Survey: Manual and Interpretation Guide*. Boston: The Health Institute, New England Medical Center, 1993.
17 Ware JE Jr, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36): conceptual framework and item selection. *Med Care* 1992; **30**: 473–83.
18 SAS Institute, Inc. *SAS System for Windows*. Cary, NC: SAS Institute, Inc.
19 Nunnaly J. *Psychometric theory*. New York: McGraw-Hill, 1978.
20 Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987; **40**: 171–8.

Paper received June 2006, accepted June 2006

