

Validation of the FACIT-fatigue subscale, selected items from FACT-thrombocytopenia, and the SF-36v2 in patients with chronic immune thrombocytopenia

James Signorovitch · Andres Brainsky ·
Kelly M. Grotzinger

Accepted: 4 April 2011 / Published online: 1 May 2011
© Springer Science+Business Media B.V. 2011

Abstract

Purpose To assess the validity and reliability of the fatigue subscale of the Functional Assessment of Chronic Illness Therapy (FACIT-F), a 6-item subset from the thrombocytopenia subscale of the Functional Assessment of Cancer Therapy (FACT-Th6) and the Short Form-36 Version 2 (SF-36v2) in 2 clinical trials of the thrombopoietin receptor agonist eltrombopag in chronic immune thrombocytopenia (ITP) patients.

Methods In the 6-month, RANdomized placebo-controlled ITP Study with Eltrombopag (RAISE; $n = 197$), the FACIT-F, FACT-Th6, and SF-36v2 were administered at baseline, day 43, weeks 14 and 26, or early withdrawal. In the ongoing open-label extension study, Eltrombopag EXTENDED Dosing Study (EXTEND; $n = 154$), measures were administered at baseline, at the beginning of each stage, and at permanent discontinuation of study medication.

Results FACIT-F, FACT-Th6, and SF-36v2 demonstrated acceptable internal consistency reliability (i.e., all Cronbach's alphas >0.70) and test-retest reliability (all intraclass correlation coefficients >0.70). Construct

validity was supported by moderate ($0.35 < r < 0.50$) to strong ($r > 0.50$) inter-measure correlations for baseline and change scores. A small to medium magnitude of effect was captured by the FACIT-F and FACT-Th6 among patients who experienced sustained platelet responses.

Conclusions Results provide support for the validity, reliability, and responsiveness of the FACIT-F, FACT-Th6, and SF-36v2 in chronic ITP patients.

Keywords Chronic ITP · FACIT-F · FACT-thrombocytopenia subscale · Fatigue · SF-36v2 · HRQoL · Thrombocytopenia · Validation studies

Abbreviations

ANOVA	Analysis of variance
EXTEND	Eltrombopag EXTENDED dosing study
FACIT-F	Fatigue subscale of the Functional Assessment of Chronic Illness Therapy
FACT-Th	Functional Assessment of Cancer Therapy-Thrombocytopenia
HRQoL	Health-related quality of life
ICC	Intraclass correlation coefficient
ITP	Immune thrombocytopenia
LOTA	Last on-treatment assessment
MCS	Mental component summary
MEI-SF	Motivation and Energy Inventory-Short Form
MID	Minimally important difference
PCS	Physical component summary
RAISE	RANdomized placebo-controlled ITP Study with Eltrombopag
SD	Standard deviation
SF-36v2	Medical Outcomes Study Short Form-36 version 2
WHO	World Health Organization

Trial registration numbers: EXTEND NCT00351468, RAISE NCT00370331.

J. Signorovitch
Analysis Group, Inc., Huntington Ave, Boston, MA, USA

A. Brainsky · K. M. Grotzinger (✉)
GlaxoSmithKline, Global Health Outcomes, Oncology,
S Collegeville Rd, Collegeville, PA, USA
e-mail: kelly.m.grotzinger@gsk.com

Background

Chronic immune thrombocytopenia (ITP) is thought to be the result of autoimmune impairment by antiplatelet antibodies, resulting in increased platelet destruction and reduced platelet production [1–3]. Chronically low platelet levels can result in bleeding into the skin and mucous membranes after minor trauma and in rare complications such as intracranial hemorrhage [3]. Chronic ITP typically occurs in adults aged 18–40 years [3]. Prevalence estimates based on a nationwide managed care database suggest that 52,700 US adults may have chronic ITP [4]. The goal of treatment is to increase and maintain platelet counts within a range that reduces the risk of bleeding and improves patient well-being without serious adverse events. Treatments include corticosteroids, IVIg, rituximab, thrombopoietin receptor agonists, platelet transfusions, and splenectomy [3].

Chronic ITP negatively affects health-related quality of life (HRQoL) due to fatigue, bleeding, bruising, and petechiae, reductions in mental and emotional health and limitations in lifestyle choices for work and leisure activities [5, 6]. Fatigue is common and can impair overall QoL and the ability to complete normal daily activities [6]. Chronic ITP can also be associated with anxiety and depression. Various side effects resulting from ITP therapies can also negatively affect HRQoL.

Despite the effects of chronic ITP and its treatments on HRQoL, there are no widely used, uniformly accepted, or validated instruments for the assessment of HRQoL in patients with this disease. However, several questionnaires assess fatigue and other HRQoL effects in patients with cancer or chronic illnesses, including the fatigue subscale of the Functional Assessment of Chronic Illness Therapy (FACIT-F) and the Functional Assessment of Cancer Therapy-Thrombocytopenia (FACT-Th). The Medical Outcomes Study Short Form-36 version 2 (SF-36v2) is a non-disease-specific questionnaire frequently used to assess multiple domains of HRQoL in a variety of disease states. These instruments have not been validated in patients with chronic ITP.

Eltrombopag is an oral, non-peptide, thrombopoietin receptor agonist that increases platelet production and reduces the risk of bleeding in patients with chronic ITP [7]. The present analysis uses data from 2 clinical studies of eltrombopag (the RANdomized placebo-controlled ITP Study with Eltrombopag [RAISE] and the Eltrombopag EXTENDED Dosing Study [EXTEND]) to assess the validity and reliability of the FACIT-F, an extract from the FACT-Th and the SF-36v2 in patients with chronic ITP. Although originally developed for use in cancer, the FACIT-F has been validated for use in selected chronic

diseases characterized by fatigue [8, 9]. The full 18-item FACT-Th questionnaire was designed based on input from patients with cancer. It contains items specific to cancer treatment and has been validated for use in patients with cancer for whom thrombocytopenia is a particular problem or side effect of treatment [10]. A subset of items from the FACT-Th was therefore chosen for assessment in RAISE and EXTEND to address concerns of ITP patients that may not be well-represented by the SF-36 and FACIT-F. In particular, focus groups conducted with ITP patients prior to the design of RAISE and EXTEND identified fatigue, psychological distress, and anxiety associated with the risk of bleeding as factors that would be impacted by effective ITP treatment. These findings are consistent with concerns of chronic ITP patients identified in more recent focus groups [6]. In the series of focus groups prior to RAISE and EXTEND, patients with chronic ITP also reviewed a collection of standard questionnaires available at that time, including the SF-36v2 and FACIT-F, and suggested that additional questions were needed to assess fears and worries associated with the risk of bleeding. A subset of 6 items from the FACT-Th was therefore selected to capture worry about bleeding, avoidance of activities due to concerns about bleeding, and frustration due to inability to carry out usual activities. Thus, the FACT-Th6 extract used in the present study consists of these 6 items. The FACT-Th6 is not an independent questionnaire or formal subscale. The SF-36 is a commonly used instrument for measuring general health status that consists of 36 items measuring 8 health domains: physical function, physical role, bodily pain, general health, vitality, social function, emotional role, and mental health. Domain scores can be further combined into summary mental and physical component scores (MCS-36 and PCS-36, respectively) [11]. The objective of this study was to assess the validity and reliability of the FACIT-F subscale, the 6 items from the FACT-Th, and the SF-36v2 in patients with chronic ITP.

Methods

Trial designs and patient populations

The RAISE study was a 6-month, phase 3, randomized, double-blind, placebo-controlled study of efficacy, safety, and tolerability of eltrombopag in patients previously treated for chronic ITP [12]. The primary endpoint of the RAISE study was the odds of achieving target platelet counts ($\geq 50,000$ and $\leq 400,000/\mu\text{L}$) during the 6-month treatment period. Patients were randomly assigned to treatment with eltrombopag ($n = 135$) or placebo ($n = 62$); individualized dose adjustments were made

during the study based on platelet counts. Platelet counts and bleeding assessments were performed weekly for at least the first 6 weeks and until a stable platelet count was achieved on a stable dose of medication, and then every 4 weeks. The FACIT-F, FACT-Th6, and SF-36v2 were completed at baseline, day 43, week 14, week 26, or at early withdrawal from study.

EXTEND is an ongoing open-label extension study to evaluate the long-term safety and efficacy of eltrombopag in patients who had previously participated in an eltrombopag study (including RAISE) [13]. The primary study endpoints are safety parameters. In Stage 1, eltrombopag is initiated at 50 mg/day; if necessary, the dose is adjusted to raise the platelet levels $\geq 100,000/\mu\text{L}$. In Stage 2, concomitant ITP medication use is minimized while maintaining platelet counts $\geq 50,000/\mu\text{L}$. In Stage 3, the minimal eltrombopag dose that maintains platelet counts $\geq 50,000/\mu\text{L}$ with no or reduced concomitant ITP medication use is determined. In Stage 4, having identified the minimal beneficial dose, long-term safety and efficacy of eltrombopag are assessed. The FACIT-F, FACT-Th6, and SF-36v2 assessments are completed at the beginning of each stage, at least every 3–6 months during each stage, prior to any new intervention or change in therapy, and upon withdrawal from the study. This study uses EXTEND data available through January 7, 2008, on 154 patients. For some analyses requiring extended exposure, fewer than 154 patients were evaluable because of the staggered enrollment into the trial.

RAISE and EXTEND also included an additional HRQoL evaluation, the Motivation and Energy Inventory-Short Form (MEI-SF), which consists of 18 questions about recent difficulties with motivation and energy [14]. In addition, regular prospective standardized evaluation of bleeding was conducted using the World Health Organization (WHO) Bleeding Scale, a single, 5-point clinician determination of bleeding severity.

HRQoL measurements

The FACIT-F is a validated symptom-specific subscale of the FACIT instrument containing 13 questions related to fatigue [8, 9]. Patients rate the frequency of a fatigue-related symptom or activity-related consequence of fatigue over the past 7 days. Each question is rated on a scale of 0 (“Not at all”) to 4 (“Very much”). All items except 7 and 8 were reverse-scored, resulting in a total possible subscale score of 0–52, with higher scores representing better HRQoL. Assessments missing 6 or more item responses were not used in the analysis. Otherwise, responses for missing items were imputed using the arithmetic average of non-missing responses.

The FACT-Th, an 18-item subscale of the FACT measuring various thrombocytopenia-related problems and concerns, was developed and validated for use in patients with cancer with thrombocytopenia [10]. The present analysis is based on the FACT-Th6, which includes 6 items selected by patients with chronic ITP. FACT-Th uses the same Likert scale as the FACIT-F, with patients rating their degree of concern in the past 7 days. The 6 selected items pertain to ability to do usual activities, worry about problems with bleeding or bruising, worry about the possibility of serious bleeding, avoidance of physical or social activity because of concern with bleeding or bruising and frustration due to the inability to carry out usual activities. All items except item 1 were reverse-scored, resulting in a range of possible scores of 0–24, with higher scores representing better HRQoL. For FACT-Th6 assessments with ≤ 3 missing responses, available responses were averaged to impute missing responses; assessments with ≥ 4 missing responses were not used in the analysis.

For the SF-36v2, all domain and summary scores were normalized (0–100). In addition, the physical component summary (PCS) and the mental component summary (MCS) scores of the SF-36v2 were further transformed to enable direct comparison with normative values derived from the 1998 US census (mean, 50; standard deviation 10) [15].

Statistical analysis

Data from the intent-to-treat populations of RAISE and EXTEND were analyzed separately and included only patients with a baseline score and at least 1 non-baseline, on-treatment score.

Common factor analysis was conducted to determine the extent to which the FACIT-F and FACT-Th6 instruments each assess a single underlying construct. Baseline values from RAISE and EXTEND were used to generate eigenvalues by spectral decomposition of item responses across patients. The number of underlying component constructs was determined by the number of eigenvalues greater than 1 [16–18]. Final communality estimates were used to assess the contribution of individual items to the measure of the underlying construct identified in the eigenvalue analysis.

Internal consistency was assessed using inter-item correlations, item-to-total score correlations, and Cronbach's alpha. Correlation coefficients of 0.20 or higher are generally accepted as indicating sufficient internal consistency [19]. When assessing internal consistency using Cronbach's alpha, values of 0.70 or higher at the group level are generally considered acceptable [20].

Test–retest reliability was assessed using a 2-way random effects intraclass correlation coefficient (ICC) [21,

27]. ICCs were calculated for all patients using assessments from consecutive pairs of visits identified as corresponding to the minimal absolute percent change in platelet counts, providing 2 repeated measures per patient. For sensitivity analyses, ICCs were also calculated from a subset of clinically stable patients using assessments from 2 consecutive visits in which platelet counts changed by 15% or less.

Construct validity was assessed by testing hypotheses about relationships between the FACIT-F, the FACT-Th6, or the SF-36v2 and other HRQoL instruments and with clinical outcomes. Pearson correlations were assessed for all pairwise comparisons of the FACIT-F; the FACT-Th6; and the SF-36v2 PCS, MCS, physical function (SF-36 PF), and vitality (SF-36 VT) domains; as well as the MEI-SF summary score. We hypothesized that related measures would have moderate ($r = 0.35\text{--}0.49$) to strong ($r \geq 0.50$) correlations to support their convergent construct validity (FACIT-F; FACT-Th6; PCS-36; SF-36 vitality; MEI-SF).

Responsiveness of the measures was assessed by analyzing differences in instrument scores between baseline and the last on-treatment assessment (LOTA) among patients with a platelet count response (responders). Responsiveness was calculated and interpreted using standardized response means (SRMs), a variant of the effect size statistics (D/SD), in which D = mean score change of interest (i.e., mean change from baseline among patients with a platelet count response), SD = standard deviation (SD) of change scores. In general, small, medium, and large effect sizes are indicated by values of 0.20–0.49, 0.50–0.79, and 0.80, respectively [22, 23]. We hypothesized that change scores of responders would be small to medium in magnitude, while mean change scores in non-responders would be trivial ($SRM < 0.20$).

Responders were defined as patients who had either: (1) a doubling of platelet count from baseline or (2) an increase in platelet counts $>50,000/\mu\text{L}$ (minimum treatment target) in patients whose baseline platelet counts were $<50,000/\mu\text{L}$. Confidence intervals were calculated for the anchor-based estimates using generalized estimating equations [24]. To facilitate comparison between the anchor-based estimates from RAISE and EXTEND, analyses of EXTEND data were conducted for all patients and separately for patients with baseline platelet counts $<30,000/\mu\text{L}$.

Results

Baseline demographics, clinical characteristics, and HRQoL scores are shown in Table 1. The mean FACIT-F score at baseline was 36.1 ($SD = 11.3$) in RAISE, below the reference value of 43.6 ($SD = 9.4$) reported for the general US population [25]. Mean scores for the SF-36v2

MCS and PCS in RAISE were below but within 1 SD of the US population standardized mean, as were mean scores for 7 of the 8 domains [11]. Item non-responses were observed for less than 2% of the items on the PRO measures except SF-36v2 at baseline (4.5%), and no scale had more than 2 missing items except one assessment of the SF-36v2 at baseline.

Common factor analysis of the FACIT-F subscale and the FACT-Th6 was conducted using baseline values from both studies to assess the degree to which the 2 instruments measure single constructs or factors. When data from the RAISE or EXTEND studies were used to calculate eigenvalues for each instrument, both instruments contained only 1 eigenvalue >1 , indicating that each instrument measures a single construct. All items were found to contribute adequately to the measured construct. In both studies, communality estimates ranged from 0.28 to 0.81 for items in the FACIT-F and, with one exception, from 0.24 to 0.73 for items in the FACT-Th6.

Internal consistency of each instrument was demonstrated by generally moderate to strong inter-item correlations in both RAISE and EXTEND. Inter-item correlations of the FACT-Th6 ranged from 0.20 to 0.85, with items 2, 3, and 4 having correlations with item 1 that were below 0.4, which suggested they did not fit the scale well. FACIT-F item-to-total score correlations ranged from 0.50 to 0.87 in the 2 studies (Table 3). All SF-36 item-to-domain score correlations exceeded 0.20.

In both studies, values for Cronbach's alpha supported acceptable levels of internal consistency; these values were ≥ 0.9 for the FACIT-F and ≥ 0.8 for the FACT-Th6. Cronbach's alpha values for SF-36v2 domains were ≥ 0.75 for all items and if each item was deleted from scale. In RAISE, Cronbach's alpha ranged from 0.75 to 0.94 at baseline and 0.83 to 0.95 at the last assessment, and in EXTEND, from 0.78 to 0.94 at baseline and 0.79 to 0.96 at the last assessment.

ICCs for test–retest reliability evaluation in clinically stable patients ranged from 0.72 to 0.78 in the FACIT-F and FACT-Th6 instruments, respectively. For the SF36v2, ICCs exceeded 0.7 in RAISE for the physical function, general health, and vitality domains ($n = 50\text{--}55$), and in EXTEND for all domains except bodily pain and emotional role ($n = 126\text{--}132$). The source of data for clinically stable patients for ICC estimation was based on 2 consecutive scores for time points when platelet counts were most similar for each patient (mean of 42 days for RAISE and 39–43 days in EXTEND). However, this analysis does not impose any absolute degree of similarity between studied pairs of visits. To further focus on visits between which patients were clinically stable, we also analyzed the subgroup of patients with a $\leq 15\%$ change in platelet counts between consecutive visits.

Table 1 Patient demographics and baseline clinical characteristics

Characteristic	RAISE		EXTEND
	Placebo <i>n</i> = 62	Eltrombopag <i>n</i> = 135	<i>n</i> = 207 ^a
Median age (min–max), y	52.5 (18–77)	47.0 (18–85)	50.0 (18–86)
Sex, <i>n</i> (%)			
Female	43 (69)	93 (69)	138 (67)
Ethnicity, <i>n</i> (%)			
White	44 (71)	101 (75)	162 (78)
African-American	1 (2)	1 (1)	1 (<1)
Asian	13 (21)	21 (15)	31 (15)
Other	4 (6)	12 (9)	13 (6)
Baseline clinical characteristics, <i>n</i> (%)			
Baseline platelet count ≤ 15,000/ μ L	30 (48)	67 (50)	94 (45)
Concomitant ITP medication	31 (50)	63 (47)	69 (33)
Splenectomy	21 (34)	50 (37)	82 (40)
≥3 prior ITP therapies	32 (52)	75 (56)	119 (57)
Median platelet counts (min–max), $\times 10^3/\mu$ L	16 (2–87)	16 (0–78)	18
Baseline HRQoL assessments, %	95%		100%
Mean baseline scores			
FACIT-F	36.3 ± 9.0	36.0 ± 12.2	36.1 ± 11.9
FACT-Th6	14.8 ± 5.8	13.5 ± 5.8	14.4 ± 6.3
SF-36v2 domain scores (mean ± SD)			
Physical function	75.0 ± 21.7	73.1 ± 26.8	73.0 ± 22.8
Physical role	64.5 ± 26.7	64.5 ± 30.5	69.1 ± 26.3
Bodily pain	70.0 ± 23.2	75.2 ± 27.8	73.9 ± 24.9
General health	53.7 ± 21.8	56.0 ± 21.3	52.1 ± 21.3
Vitality	56.7 ± 20.2	55.1 ± 26.3	54.6 ± 23.7
Social function	76.1 ± 21.7	72.7 ± 28.3	75.6 ± 23.9
Emotional role	73.4 ± 25.4	69.1 ± 30.9	76.4 ± 24.0
Mental health	70.3 ± 18.7	68.0 ± 21.0	68.8 ± 19.6
SF-36v2 summary measures (mean ± SD)			
PCS	45.6 ± 8.3	46.9 ± 9.7	46.1 ± 8.1
MCS	46.4 ± 10.1	44.3 ± 12.6	46.1 ± 11.3

^a Only data from 154 patients with both baseline HRQoL and at least one on-treatment HRQoL assessment were included in the analysis

These visits were separated by a mean of 49–52 days for RAISE and 45–50 days in EXTEND. ICCs in clinically stable patients were 0.79 for the FACIT-F subscale, 0.83 for the FACT-Th6, and ≥ 0.72 for all SF36-v2 domains and summary measures except social function and emotional role.

Construct validity of the FACIT-F, FACT-Th6, MEI-SF, and SF-36v2 was supported in both RAISE and EXTEND by moderate to strong score correlations between scores at baseline (Table 2). Moderate to strong correlations between the change scores of the measures were also observed (Table 3).

To further evaluate longitudinal construct validity of measures, we stratified patients into responders (i.e., sustained increases in platelet counts) or non-responders, and compared the change score on each measure between groups based on magnitude of effect. Mean change scores on the FACIT-F were statistically significantly larger for responders, with a small SRM in both the RAISE (SRM = 0.34) and EXTEND (SRM = 0.40) (Table 4). For the FACT-Th6, responders also demonstrated significantly more change (P values <0.01), with medium effect sizes found in both the RAISE (SRM = 0.57) and EXTEND (SRM = 0.56) studies. The SRM for non-responders was trivial (SRM < 0.20).

Table 2 Correlation between baseline scores of PRO measures

Measure #1	Measure #2	Correlation (<i>r</i>)	
		RAISE (<i>n</i> = 189)	Extend (<i>n</i> = 154)
SF-36 PF	MEI-SF	0.54	0.46
SF-36 VT	MEI-SF	0.80	0.78
PCS-36	MEI-SF	0.52	0.42
SF-36 PF	FACIT-F	0.60	0.51
SF-36 VT	FACIT-F	0.82	0.85
PCS-36	FACIT-F	0.66	0.56
SF-36 PF	FACT-Th6	0.50	0.47
SF-36 VT	FACT-Th6	0.52	0.52
PCS-36	FACT-Th6	0.50	0.45
MEI-SF	FACIT-F	0.80	0.81
MEI-SF	FACT-Th6	0.58	0.59
FACIT-Fatigue Scale	FACT-Th6	0.62	0.61

Table 3 Correlation between change scores of PRO measures

Measure #1	Instrument	Correlation (<i>r</i>)	
		RAISE (<i>n</i> = 185–186)	Extend (<i>n</i> = 150–153)
SF-36 PF	MEI-SF	0.53	0.39
SF-36 VT	MEI-SF	0.55	0.58
PCS-36	MEI-SF	0.48	0.45
SF-36 PF	FACIT-F	0.56	0.39
SF-36 VT	FACIT-F	0.63	0.62
PCS-36	FACIT-F	0.51	0.35
SF-36 PF	FACT-Th6	0.35	0.33
SF-36 VT	FACT-Th6	0.32	0.39
PCS-36	FACT-Th6	0.44	0.33
MEI-SF	FACIT-F	0.74	0.76
MEI-SF	FACT-Th6	0.43	0.50
FACIT-Fatigue Scale	FACT-Th6	0.50	0.54

Discussion

The results of this study support the validity and reliability of the FACIT-F subscale, FACT-Th6, and the SF-36v2 in patients with chronic ITP for the purpose of assessing fatigue, concerns about bleeding, and general health status, respectively.

In both RAISE and EXTEND studies, internal consistency reliability was found to be generally acceptable when evaluated for baseline scores and LOTA, i.e., Cronbach's alpha >0.70, moderate to strong correlations between most items and total scores. High Cronbach's alpha values of

0.94 and 0.95 on the FACIT-F suggested some item redundancy on that scale but these values are consistent with reliability reported by the developers [8]. Test–retest reliability was acceptable for all of the measures (ICCs > 0.70).

The construct validity of the 3 instruments in patients with chronic ITP was supported in both studies, with moderate to strong correlations between the FACIT-F subscale total score and the SF-36v2 PCS and physical functioning and vitality domain scales, as well as between the FACIT-F total score and the MEI-SF and FACT-Th6 scales. Longitudinal construct validity was supported by moderate to strong correlations between change scores for the FACIT-F, FACT-Th6, and SF-36v2. In addition, non-trivial change scores were observed in patients who sustained increases in platelet counts, suggestive that the FACIT-F and FACT-Th6 are responsive to clinically meaningful change. While the SF-36v2 demonstrated a statistically significant difference between responders and non-responders, it was less responsive than the disease-specific measures of fatigue in terms of ability to capture change. Less sensitivity to change is a common criticism of generic measures compared with disease-specific measures [26]. Analysis of the change scores suggested a trend that the FACT-Th6 may be slightly more sensitive to change than the FACIT-F in measuring issues of fatigue in patients with ITP.

The results of the present study lend support to the utility of FACIT-F and FACT-Th6 across patient populations in which fatigue is an issue [8–10]. These instruments incorporate issues and concerns related to fatigue in patients with chronic ITP. Among the strengths of the present evaluation is the inclusion of a large sample of patients representative of those with chronic ITP who have failed first-line treatments.

There were several study limitations. The same data used to assess the results of clinical trials were used to test the psychometric performance of these scales. There was a lack of responses that indicated extreme fatigue at baseline and on-treatment assessments. There was a large time interval between assessments for the test–retest reliability analysis, and it was unclear whether patients were really stable in fatigue. Instruments for assessing validity were not previously validated in patients with chronic ITP, so we examined the construct validity using multiple methods and against a variety of measures. As it is not a patient-reported criterion, we acknowledge that platelet count has limitations as a criterion for fatigue and as a basis for examining responsiveness. As EXTEND was open-label, patient knowledge of their platelet count might introduce bias, but the consistency of results between RAISE and EXTEND alleviate this concern.

Table 4 Differences in PRO score changes between responders and non-responders

PRO measure	Changes in PRO score			Changes in PRO score		
	RAISE			EXTEND		
	Mean (SD)	SRM	<i>P</i> value Response vs. no response	Mean (SD)	SRM	<i>P</i> value Response vs. no response
FACIT-F subscale summary score						
Responders	3.39 (9.92)	0.34	0.0460*	4.10 (10.30)	0.40	0.0021 [†]
Non-responders	0.70 (8.25)	0.08		−0.84 (7.74)	0.11	
FACT-Th6 subscale summary score						
Responders	3.19 (5.62)	0.57	0.0001 [‡]	2.75 (4.92)	0.56	0.0092 [‡]
Non-responders	0.31 (4.46)	0.07		0.59 (4.53)	0.13	
SF-36v2 summary scores						
PCS						
Responders	2.42		0.04			
Non-responders	0.16					
MCS						
Responders	2.44		0.06			
Non-responders	−0.41					

MCS mental component summary, *PCS* physical component summary, *PRO* patient-reported outcome, *SD* standard deviation, *SRM* standardized response mean

* $P < 0.05$; [†] $P < 0.01$; [‡] $P < 0.001$

Conclusions

The results of this study provide evidence that generally supports the internal consistency, test–retest reliability, and construct validity of the SF-36, FACIT-F, and FACT-Th6 in patients with chronic ITP. The FACIT-F subscale detected small to moderate improvements in HRQoL associated with improved clinical status in patients with chronic ITP. Future research should further investigate the properties of these scales in non-clinical trials populations of patients with chronic ITP.

Acknowledgments Funding for this study was provided by GlaxoSmithKline. All listed authors meet the criteria for authorship set forth by the International Committee for Medical Journal Editors. The authors wish to acknowledge the following individuals for their contributions and critical review during the development of this manuscript: Lisa Breck, Steve Collins, and AOI Communications, L.P., for editorial assistance, and Kimberly Marino for critical review.

References

- Ballem, P. J., Segal, G. M., Stratton, J. R., Gernsheimer, T., Adamson, J. W., & Slichter, S. J. (1987). Mechanisms of thrombocytopenia in chronic autoimmune thrombocytopenic purpura. Evidence of both impaired platelet production and increased platelet clearance. *The Journal of Clinical Investigation*, *80*, 33–40.
- Chang, M., Nakagawa, P. A., Williams, S. A., Schwartz, M. R., Imfeld, K. L., Buzby, J. S., et al. (2003). Immune thrombocytopenic purpura (ITP) plasma and purified ITP monoclonal auto-antibodies inhibit megakaryocytopoiesis in vitro. *Blood*, *102*, 887–895.
- Cines, D. B., & Blanchette, V. S. (2002). Immune thrombocytopenic purpura. *New England Journal of Medicine*, *346*, 995–1008.
- Feudjo-Tepie, M. A., Robinson, N. J., & Bennett, D. (2008). Prevalence of diagnosed chronic immune thrombocytopenic purpura in the us: Analysis of a large us claim database: A rebuttal. *Journal of Thrombosis Haemostasis*, *6*, 711–712, author reply 713.
- Mathias, S. D., Bussel, J. B., George, J. N., McMillan, R., Okano, G. J., & Nichol, J. L. (2007). A disease-specific measure of health-related quality of life for use in adults with immune thrombocytopenic purpura: Its development and validation. *Health and Quality of Life Outcomes*, *5*, 11.
- Mathias, S. D., Gao, S. K., Miller, K. L., Cella, D., Snyder, C., Turner, R., et al. (2008). Impact of chronic immune thrombocytopenic purpura (ITP) on health-related quality of life: A conceptual model starting with the patient perspective. *Health and Quality of Life Outcomes*, *6*, 13. Available at <http://www.hqlo.com/content/16/1/13>.
- Bussel, J. B., Cheng, G., Saleh, M. N., Psaila, B., Kovaleva, L., Meddeb, B., et al. (2007). Eltrombopag for the treatment of chronic idiopathic thrombocytopenic purpura. *New England Journal of Medicine*, *357*, 2237–2247.
- Cella, D., Eton, D. T., Lai, J. S., Peterman, A. H., & Merkel, D. E. (2002). Combining anchor and distribution-based methods to derive minimal clinically important differences on the functional assessment of cancer therapy (FACT) anemia and fatigue scales. *Journal of Pain and Symptom Management*, *24*, 547–561.

9. Webster, K., Cella, D., & Yost, K. (2003). The functional assessment of chronic illness therapy (FACIT) measurement system: Properties, applications, and interpretation. *Health and Quality of Life Outcomes*, *1*, 79.
10. Cella, D., Beaumont, J. L., Webster, K. A., Lai, J. S., & Elting, L. (2006). Measuring the concerns of cancer patients with low platelet counts: The functional assessment of cancer therapy–thrombocytopenia (FACT-Th) questionnaire. *Supportive Care in Cancer*, *14*, 1220–1231.
11. Ware, J. E., Kosinski, M., Bjorner, J. B., Turner-Bowker, D. M., Gandek, B., & Mariush, M. E. (2007). *User's manual for the sf-36v2 health survey* (2nd ed.). Lincoln, RI: Quality Metric Incorporated.
12. Cheng, G., Saleh, M. N., Marcher, C., Vasey, S., Mayer, B., Aivado, M., et al. (2010). Eltrombopag for the long-term management of chronic idiopathic thrombocytopenic purpura: A 6-month, randomized, phase iii study (RAISE). *Lancet*. Epub ahead of print.
13. Saleh, M. N., Bussel, J. B., Cheng, G., Meddeb, B., Mayer, B., Bailey, C., et al. (2009). Long-term treatment of chronic immune thrombocytopenic purpura with oral eltrombopag: Results from the extend study. *Blood*, *114*, Abstract 682.
14. Fehnel, S. E., Bann, C. M., Hogue, S. L., Kwong, W. J., & Mahajan, S. S. (2004). The development and psychometric evaluation of the motivation and energy inventory (MEI). *Quality of Life Research*, *13*, 1321–1336.
15. Marra, C. A., Woolcott, J. C., Kopec, J. A., Shojania, K., Offer, R., Brazier, J. E., et al. (2005). A comparison of generic, indirect utility measures (the hui2, hui3, sf-6d, and the eq-5d) and disease-specific instruments (the raqol and the haq) in rheumatoid arthritis. *Social Science and Medicine*, *60*, 1571–1582.
16. Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika*, *19*, 149–161.
17. Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151.
18. Kaiser, H. F. (1961). A note on Guttman's lower bound for the number of common factors. *British Journal of Mathematical and Statistical Psychology*, *14*, 1.
19. Kline, P. (1993). *Handbook of psychological testing*. London: Routledge.
20. Bland, J. M., & Altman, D. G. (1997). Cronbach's alpha. *British Medical Journal*, *314*, 572.
21. Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in addressing rater reliability. *Psychological Bulletin*, *86*, 420–428.
22. Cohen, J. (1977). *Statistical power analysis for the behavioral science* (revised ed.) New York: Academic Press.
23. Moretti, C., Viola, S., Pistorio, A., Magni-Manzoni, S., Ruperto, N., Martini, A., et al. (2005). Relative responsiveness of condition specific and generic health status measures in juvenile idiopathic arthritis. *Annals of the Rheumatic Diseases*, *64*, 257–261.
24. Liang, K.-Y., & Seger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*, 13–22.
25. Cella, D., Lai, J. S., Chang, C. H., Peterman, A., & Slavin, M. (2002). Fatigue in cancer patients compared with fatigue in the general united states population. *Cancer*, *94*, 528–538.
26. Guyatt, G. H., King, D. R., Feeny, D. H., Stubbings, D., & Goldstein, R. S. (1999). Generic and specific measurement of health-related quality of life in a clinical trial of respiratory rehabilitation. *Journal of Clinical Epidemiology*, *52*(3), 187–92. PubMed PMID: 10210235.
27. McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*, 30–46.