



Psychometric Evaluation of ITP Life Quality Index (ILQI) in a Global Survey of Patients with Immune Thrombocytopenia

Ricardo Viana · Denise D'Alessio · Laura Grant · Nichola Cooper · Donald Arnold · Mervyn Morgan · Drew Provan · Adam Cuker · Quentin A. Hill · Yoshiaki Tomiyama · Waleed Ghanima

Received: June 10, 2021 / Accepted: September 23, 2021 / Published online: October 27, 2021
© The Author(s) 2021

ABSTRACT

Introduction: Immune thrombocytopenia (ITP) is an autoimmune disorder caused by immunologic destruction of otherwise normal platelets. Patients and physicians differ in their views pertaining to the limitations imposed on patients' daily lives by ITP and its treatment.

Poor understanding of ITP symptoms can result in misdiagnosis and complex treatment patterns, and affect patient health-related quality of life (HRQoL). The ITP Life Quality Index (ILQI) is a 10-item patient-reported outcome measure developed for clinical practice to aid discussions between patients and physicians. This research aimed to validate the

R. Viana
Novartis Pharma AG, Basel, Switzerland

D. D'Alessio
Novartis Pharmaceuticals Corporation, East Hanover, NJ, USA

L. Grant (✉)
Adelphi Values Ltd, Adelphi Mill, Grimshaw Lane, Bollington, Cheshire SK10 5JB, UK
e-mail: laura.grant@adelphivalues.com

N. Cooper
Department of Haematology, Hammersmith Hospital, Imperial College London, London, UK

D. Arnold
Department of Medicine, McMaster Centre for Transfusion Research, McMaster University, Hamilton, ON, Canada

M. Morgan
ITP Support Association, Bolnhurst, UK

D. Provan
Academic Haematology Unit, Blizard Institute, Barts and The School of Medicine and Dentistry, London, UK

A. Cuker
Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

A. Cuker
Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

Q. A. Hill
Leeds Teaching Hospital NHS Trust, St James' University Hospital, Leeds, UK

Y. Tomiyama
Department of Blood Transfusion, Osaka University Hospital, Suita, Japan

W. Ghanima
Ostfold Hospital Trust, Gralum, Norway

W. Ghanima
Department of Hematology, Institute of Clinical Medicine, University of Oslo, Oslo, Norway

psychometric properties of the ILQI using data collected in the ITP World Impact Survey (I-WISH).

Methods: I-WISH data containing responses to the ILQI from 1507 patients with ITP across 13 countries worldwide was subject to psychometric analysis to evaluate the structure, reliability and validity of the ILQI and assess scoring cut-offs.

Results: The ILQI has an overarching unidimensional structure, supporting a total score including all 10 items. Reliability was supported (Cronbach's $\alpha = 0.90$). ILQI scores monotonically increased with ITP severity. ILQI scores correlated with measures of fatigue and emotional well-being, supporting construct validity. Differential item functioning (DIF) analyses showed that ILQI item responses were interpreted similarly between the USA and other Western countries. It was suggested that previous clinical cut-off score of 20 for "impaired HRQoL" was reduced to 17 and a cut-off of 23–25 (rather than 30) was suggested to assess "significantly impaired HRQoL".

Conclusion: The validity and reliability of the ILQI to assess HRQoL of patients with ITP is supported. The revised cut-off scores for the ILQI will aid patient-centric decision-making.

Keywords: Health related quality of life (HRQoL); Immune thrombocytopenia (ITP); ITP Life Quality Index (ILQI); ITP World Impact Survey (I-WISH); Psychometric evaluation; Tool for clinical practice

Key Summary Points

Why carry out this study?

To assess the psychometric properties of the ILQI and confirm the scoring cut-offs, in order to evaluate the reliability and validity of scores and demonstrate that ILQI is appropriate for measuring the impact of ITP on HRQoL in clinical practice.

What did the study ask? What was the hypothesis of the study?

Hypothesis was that the ILQI had already demonstrated good content validity and developers wanted to assess the psychometric validity, to further support appropriateness for use in clinical practice.

What were the study outcomes/conclusions?

Psychometric evaluation demonstrates that the ILQI is fit for measuring the impact of ITP on quality of life.

What has been learned from the study?

ILQI is a valid and reliable PRO measure of quality of life in patients with ITP.

INTRODUCTION

Immune thrombocytopenia (ITP) is an autoimmune disorder characterized by immunologic destruction of platelets, leading to an increased risk of bleeding. ITP is considered a rare disease, diagnosed primarily by excluding other possible causes of disease, and symptoms can present with varying severities. Patients often experience misdiagnosis and complex treatment patterns, which may significantly affect health-related quality of life (HRQoL) in this patient population.

Bleeding is identified as the primary symptom experienced by patients with ITP; however, this can present across a broad range of severity, from mild bruising and mucosal bleeding to severe haemorrhage [1]. A wide range of bleeding manifestations exist and the primary goal of treatment is to prevent severe/life-threatening bleeding [2]. Consequently, first-line therapy is indicated for patients with bleeding complications and those who are at increased risk of bleeding; however, the decision to initiate therapy depends not only on platelet count and the risk of bleeding but also on other outcomes,

including quality of life and multiple lifestyle variables [3, 4]. Improving HRQoL and patient outcomes are key considerations when evaluating the impact of disease and the burden and benefits of treatments [5].

While patients and physicians largely align on overall ITP symptom burden, some differences can arise in their views pertaining to the limitations caused by ITP and/or its treatment [6, 7]. Employment, marital status, and education may influence HRQoL in patients with ITP and should be considered in clinical practice, as well as when planning and conducting studies [8]. The International Consensus Report has mentioned that severe fatigue is the most difficult ITP symptom to treat, experienced by 39–59% of adult patients, yet it is under-recognized by healthcare practitioners. Further, physicians primarily focus on platelet count whereas fatigue and symptoms related to mental health are often the more primary concerns of patients. The consensus report recommends individualization of treatment, and that improvement in patients' HRQoL should be a primary treatment goal [9].

There is currently no appropriate disease-specific prospective tool in routine clinical practice to quantify the HRQoL of adult patients with ITP. The ITP Life Quality Index (ILQI) is a 10-item patient-reported outcome measure (Fig. 1) developed for use in clinical practice to aid discussions between patients and physicians about patients' disease experience, and therefore facilitate a patient-centric approach towards treatment decision-making. The ILQI was originally developed by clinical experts in the field of ITP and content validity was confirmed by conducting qualitative interviews with 15 adult patients. The ILQI was also cognitively debriefed and items refined following qualitative analysis and additional clinical input [10, 11].

The ILQI was included in the ITP World Impact Survey (I-WISH) [12, 13], a global observational survey which collected real-world data characterising the experiences of patients with ITP [6]. This large study provided an opportunity to evaluate the psychometric properties of the ILQI in order to assess the reliability and validity of scores and

demonstrate that ILQI is appropriate for measuring the impact of ITP on HRQoL in clinical practice.

METHODS

I-WISH Data

The ILQI was administered alongside a variety of other clinical and demographic questions in the I-WISH [12], including a series of single-item questions assessing HRQoL and items adapted from the Work Productivity and Activity Impairment Questionnaire (WPAI) [14]. The global survey was completed by 1507 adult patients with ITP and 472 healthcare providers from 13 countries (in order of number of patients recruited: USA, China, UK, France, Germany, Italy, India, Canada, Turkey, Japan, Colombia, Spain and Egypt). Patients were recruited via physicians and patient support groups; all patients self-confirmed their diagnosis of ITP, confirmed they were over 18 years old and gave their consent to participate. Healthcare providers confirmed their primary speciality was either haematology or haematology-oncology, confirmed their caseload of patients with ITP, and confirmed that they had responsibility for making ITP treatment decisions. Patients and healthcare providers were provided with an online link to complete study questions.

The ILQI

The ILQI is a 10-item patient-reported outcome (PRO) measure with a recall period of 'the last month' (Fig. 1). The items generally have four response options ranging from 'never' to 'all of the time' and three items (items 1, 2 and 5) have additional response options which allow the patient to specify that the item does not apply to them or that they do not wish to answer.

The ILQI was originally developed by clinical experts in the field of ITP and was considered unidimensional, resulting in a single score. Originally a total sum score ranging between 7 and 40 was proposed, where coded responses were summed (see Fig. 1 for coding). As the total

ITP Life Quality Index (ILQI)

The aim of this questionnaire is to measure how much your ITP has affected your life **OVER THE PAST MONTH**. The aim is to try to standardise how, besides bleeding, your ITP affects your life. Please tick one box.

1. How often has your ITP impacted on your working life or studies?

- Never Sometimes More than half the time All the time
 I am not currently working/studying due to ITP
 I am not currently working/studying due to other reasons (0)

2. How often have you taken time off work or education because of your ITP?

- Never Sometimes More than half the time All the time
 I am not currently working/studying due to ITP
 I am not currently working/studying due to other reasons (0)

3. How often has your ITP impacted your ability to concentrate on everyday tasks?

- Never Sometimes More than half the time All the time

4. How often has your ITP impacted your social life?

- Never Sometimes More than half the time All the time

5. How often has your ITP impacted your sex life?

- Never Sometimes More than half the time All the time Not applicable/prefer not to say

6. How often has your ITP impacted your energy levels?

- Never Sometimes More than half the time All the time

7. How often has your ITP impacted your undertaking of daily tasks?

- Never Sometimes More than half the time All the time

8. How often has your ITP impacted your ability to support people close to you?

- Never Sometimes More than half the time All the time

9. How often has your ITP negatively impacted your hobbies?

- Never Sometimes More than half the time All the time

10. How often has your ITP negatively impacted your normal capacity to exercise?

- Never Sometimes More than half the time All the time

Response coding: Never=1, Sometimes=2, More than half of the time=3, All the time=4, I am not currently working/studying due to ITP=4, I am not currently working due to other reasons=0, not applicable/prefer not to say=0

Fig. 1 ITP Life Quality Index (ILQI)

sum score did not account for missing items, an alternative scoring procedure was tested in this analysis, using the mean of all items transformed to reflect the initial maximum total score of 40. This allows for the calculation of a total score when there are missing items; henceforth, this score is referred to as ILQI total (alternative) score.

The calculation for the ILQI total (alternative) score is as follows:

ILQI total (alternative) score = total sum score of all answered items/number of items with a score $\geq 1 \times 10$.

Prior to the psychometric validation work, a score of 20 or above was defined by the clinical experts as the cut-off for 'impaired' HRQoL and a score of 30 or above was the cut-off for 'significantly impaired' HRQoL.

Statistical Analyses

The normality of the data and individual ILQI items were assessed. Item response distribution and floor and ceiling effects were examined for each item to ensure the response scale appropriately captured the range of possible severity. Floor and ceiling effects were defined as $> 25.00\%$ of patients selecting the worst or best possible response respectively (assuming a uniform distribution across the four main responses). Underused response options (those selected by $< 5.00\%$ of patients) were also flagged for further consideration.

Differential item functioning (DIF) was conducted to assess whether patients with ITP in countries other than the USA answered each item in a similar way to the USA cohort of patients, stratified by disease severity. Uniform and non-uniform DIF was assessed using ordinal logistic regression models with item response as the dependent variable, and ILQI total (alternative) score and country indicator as covariates [15]. Note that sample sizes of at least 200 patients per group are recommended for DIF analyses [16]; therefore, this analysis was exploratory in nature due to the limited likelihood of achieving such sample sizes.

Polychoric correlation coefficients between each pair of ILQI items were reviewed to explore

dimensionality. Items with particularly high inter-item correlations (> 0.90) were considered for potential redundancy. Exploratory factor analysis (EFA) was conducted to explore the latent structure (i.e. underlying structure) of the ILQI without imposing any preconceived structure. EFA was performed for half of the analysis population, stratified on the basis of current health, age and gender and was estimated using mean- and variance-adjusted weighted least squares (WLSMV) and employing delta parameterisation. Geomin rotation was specified, which allowed for the extracted factors to be correlated. The remaining half of the population was used in a confirmatory factor analysis (CFA) and the factor structure suggested through the EFA was treated as the primary hypothesized model. Model fit was assessed by calculating the comparative fit index (CFI; > 0.95 for evidence of acceptable fit) [17], root mean square error of approximation (RMSEA; < 0.10 for acceptable fit) [18], and standardised root mean square residual (SRMR; < 0.10 for acceptable fit) [18]. The chi-square test of model fit was not assessed because of its tendency to reject the null hypothesis in large samples, even when the hypothesised model shows trivial misfit [19].

Internal consistency reliability was conducted to assess how well the ILQI items were measuring the same underlying construct (i.e. homogeneity of the items belonging to the same score). A Cronbach's alpha coefficient ≥ 0.70 was considered reliable [20] and was reassessed with each item removed in turn, to identify if items were not contributing to the reliability. The impact on missing items on the reduction in reliability was assessed through applying the Spearman–Brown prophecy formula to the Cronbach's alpha coefficient.

Convergent validity correlations were conducted to evaluate the relationship of the ILQI total (alternative) score and specific ILQI items with other measures or items. Hypothesised strong correlations (≥ 0.50) and moderate correlations (≥ 0.30 but < 0.50) are presented in Table 1.

Known-groups methods evaluated the construct validity by segregating patients into hypothetically distinct groups, according to

Table 1 Hypothesised convergent validity relationships

Known group	ILQI total (alternative) score	All ILQI items	Item 1: impacted work or study	Item 2: time off from work or study	Item 3: concentrate on every day tasks	Item 4: impact social life
Current health	≥ 0.50 (PS)	≥ 0.40 (PC)				
Fatigue	≥ 0.30 (BS)	≥ 0.30 (BS)				
Depression	–	–			≥ 0.30 (BS)	
Hide bleed signs	–	–				≥ 0.30 (PS)
Emotional well-being	≥ 0.40 (PS)	–				
Work hours missed (ITP)	≥ 0.30 (P) ^a	–	≥ 0.40 (PC) ^a	≥ 0.40 (PC) ^a		
Percentage work hours missed	≥ 0.30 (P)	–	≥ 0.40 (PC)	≥ 0.40 (PC)		
Work productivity reduction	≥ 0.30 (P)	–	≥ 0.40 (PS)	≥ 0.40 (PS)		
Daily activity reduction	≥ 0.40 (P)	≥ 0.30 (PS)				

BS bi-serial, P Pearson's *r*, PC polychoric, PS polyserial

^aExploratory: As this score is not standardized between patients, a correlation may not exist

current health (poor, moderate, good), impact on emotional well-being (low, moderate, large impact), effect on daily activity reduction (low, moderate, high impact), fatigue (yes, no) and depression (yes, no). The amount of difference in mean ILQI total (alternative) scores characterises the degree to which the score is capable of distinguishing among groups hypothesised a priori to be clinically distinct. Between-group effect sizes were calculated and interpreted as 'small' (0.20), 'medium' (0.50) and 'large' (0.80) [21].

Existing severity thresholds for ILQI were derived from clinical expert judgement at the time of original development (a score of ≥ 20 for "impaired" HRQoL and a score of ≥ 30 for "significantly impaired" HRQoL). It was therefore necessary to assess these thresholds using statistically robust methods (i.e. receiving operating characteristic [ROC] analysis). ROC analysis was used to identify the threshold on ILQI total (alternative) score which optimally discriminated between 'low', 'moderate' and 'high' severity groups according to the current health, impact on emotional well-being, and effect on daily activity reduction anchors. The threshold which had the best trade-off between

sensitivity and specificity according to the sum of squares method was selected [22]. If the area under the curve (AUC) confidence interval (CI) contains 0.50, the score may be no better than chance at predicting changes as defined by the anchor and ideally this should be > 0.70 [23, 24].

For each scoring algorithm derived as part of the analyses, the reduction in accuracy and level of bias were presented when patients were missing between one and five items using simulation of missing data (bootstrapping technique) [25].

The study was conducted in accordance with the Helsinki Declaration of 1964, and its later amendments. Survey materials and the study protocol were reviewed and approved by central institutional review boards (IRB) in both Europe and North America. All participants provided informed consent prior to completing the survey. No participant identifiable information is included in the manuscript.

RESULTS

Demographic and Clinical Characteristics (I-WISH)

The mean (\pm SD) age of patients was 46.90 ± 16.22 years and 64.70% ($n = 975$) were female (Table 2). Patients were recruited from 13 different countries. Over half of the sample were working, either full time or part-time (60.40%). A quarter of patients (25.40%) had a current platelet count of $> 100 \times 10^9/L$ and the sample included some patients with a platelet count of $< 10 \times 10^9/L$ at time of survey (5.10%). Mean time since diagnosis was 8.9 ± 10.8 years. Most of the sample (85.20%) self-reported their current health state to be good or excellent (> 4 on a 1–7 Likert scale). Further results from the I-WISH survey are previously published [6, 12, 26–28].

Descriptive Analyses

The mean ILQI total (alternative) score was 21 (SD 7.00) (Table 3) and did not deviate from statistical normality (Shapiro–Wilk $W = 0.97$). Half of the patients (49.90%) were currently experiencing fatigue and 17.90% were currently experiencing depression. Two thirds of patients (62.30%) reported hiding the signs of bleeding to some degree and almost all patients (93.30%) reported some degree of worry about platelet counts going up and down. Patients missed a mean of 6 h (SD 15.10) of work in the past 7 days. Using a 10-point NRS, patients reported the effect of ITP on their work productivity over the past 7 days as a mean of 4 (SD 2.60), with higher scores representing worse work productivity.

ILQI Item Properties

Patients used the full range of ILQI item response options. Six items displayed ceiling effects and for three of these items, the most severe end of the scale was underused (answered by fewer than 5% of respondents) (Fig. 2). The ceiling effects range from 27.80% to 40.10% of patients selecting the best possible response for

Table 2 Participant demographic and clinical characteristics

Demographic/clinical characteristics	Patients (N = 1507)
Age	
Mean (SD)	46.9 (16.22)
Range	18–90 years
Missing	$n = 1$
Gender, n (%)	
Female	975 (64.7%)
Male	532 (35.3%)
Country, n (%)	
USA	501 (33.2%)
China	286 (19.0%)
UK	120 (8.0%)
France	87 (5.8%)
Germany	82 (5.4%)
Italy	74 (4.9%)
India	65 (4.3%)
Canada	61 (4.0%)
Turkey	60 (4.0%)
Japan	56 (3.7%)
Colombia	51 (3.4%)
Spain	48 (3.2%)
Egypt	16 (1.1%)
Current employment status, n (%)	
Working full time	624 (44.6%)
Retired	267 (19.1%)
Working part time	221 (15.8%)
Homemaker	82 (5.8%)
Student	60 (4.3%)
Disability allowance	34 (2.4%)
Long-term sick leave	13 (0.9%)
Other (not working but seeking employment or not working but not seeking employment)	99 (7.1%)

Table 2 continued

Demographic/clinical characteristics	Patients (<i>N</i> = 1507)
Patient self-reported current health state, <i>n</i> (%) (7-point Likert scale, 7 = excellent health)	
1–3 (low score)	207 (14.8%)
4	306 (21.9%)
5–7 (high score)	887 (63.3%)
Current platelet counts, <i>n</i> (%)	
< 10 ($\times 10^9/L$)	77 (5.1%)
10–29 ($\times 10^9/L$)	175 (11.6%)
30–39 ($\times 10^9/L$)	145 (9.6%)
40–49 ($\times 10^9/L$)	145 (9.6%)
50–69 ($\times 10^9/L$)	195 (12.9%)
70–100 ($\times 10^9/L$)	250 (16.6%)
> 100 ($\times 10^9/L$)	383 (25.4%)
Don't know	136 (9.0%)
NS	1 (0.1%)
Time since ITP diagnosis	
Mean (SD)	8.9 years (10.75)
Range	0.1–89.7 years

an item; however, this does not necessarily indicate an issue with the performance of the items, as other factors such as disease severity of the sample need to be considered (i.e. 63.30% of this sample reported their current health state was between 5 and 7 on a 7-point Likert scale, where 7 = excellent). Item 8, assessing ability to support people close to you, had the largest ceiling effect (40.10%) and only 4.80% of patients selected the worst possible response option (i.e. ITP impacts your ability to support people close to you all of the time), suggesting the response scale may not appropriately capture the range of possible severity levels.

Table 3 PRO descriptive analysis

Characteristics	Patients
ILQI total (alternative) score	
<i>N</i>	1506
Mean (SD)	21 (7.0)
Range	10–40
Missing	1 ^a
Do you currently have fatigue? <i>n</i> (%)	
Yes	752 (49.9%)
No	755 (50.1%)
Other clinical and demographic questions	
Do you currently have depression? <i>n</i> (%)	
Yes	270 (17.9%)
No	1237 (82.1%)
Hide signs of bleeding, <i>n</i> (%)	
“I often wear long-sleeved clothing, even when it is warm, to hide the signs of bleeding”	
1—not at all	568 (37.7%)
2	216 (14.3%)
3	134 (8.9%)
4	166 (11.0%)
5	141 (9.4%)
6	131 (8.7%)
7—a great deal	151 (10.0)
Impact on emotional well-being, <i>n</i> (%)	
“I worry that my platelet count goes up and down for no apparent reason”	
1—not at all	101 (6.7%)
2	200 (13.3%)
3	214 (14.2%)
4	253 (16.8%)
5	322 (21.4%)
6	220 (14.6%)
7—a great deal	197 (13.1%)
Work hours missed, <i>n</i> (%)	

Table 3 continued

Characteristics	Patients
“During the past 7 days, how many hours did you miss from work because of problems associated with your ITP?”	
<i>N</i>	893
Mean (SD)	6 (15.1)
Range	0–168
Missing	614
Effect on work productivity, <i>n</i> (%)	
“During the past 7 days, how much did your ITP affect your productivity while you were working?”	
1 = ITP had no effect on my work, 10 = ITP completely prevented me from working	
<i>N</i>	898
Mean (SD)	4 (2.6)
Range	1–10
Missing	609

^aOf note, one patient only answered the N/A questions and not any other question, therefore a score could not be calculated using the alternate scoring procedure. Therefore there is a missing *n* = 1. A score could be calculated for the total score, although this score will not be truly accurate as the patient only answered the N/A options

Inter-item correlations were examined (Table 4). All inter-item correlations were below 0.90, suggesting no item redundancy. Lowest correlations [29] of < 0.40 were found between two specific items (item 2 and 5) with most other items of the ILQI. Both item 2: ‘how often have you taken time off work or education because of your ITP’ (*r* = 0.25–0.39) and item 5: ‘how often has ITP impacted your sex life?’ (*r* = 0.27–0.32) had an N/A or ‘prefer not to say’ option, which may have contributed to the poor correlations.

The number of patients within each country was less than the recommended sample size of 200 for DIF analyses in all countries except the USA and China (Table 2). Therefore, DIF analyses were viewed as hypothesis-generating. The

extent of uniform and non-uniform DIF was generally low when comparing the USA with Canada and other Western countries; however, items related to work or education, social life and sex life exhibited DIF between the USA and countries such as China, Turkey, Colombia and India.

Domain Structure and Scoring

The EFA factor loadings suggested that the ILQI has a unidimensional structure, supporting the creation of a total score including all 10 items. However, a two-factor solution (i.e. dividing the items into two groups of concepts which relate to each other) could be used by separating out items 1 and 2, assessing ‘work and study’. The factor loadings for the EFA sample (one- and two-factor solution) are presented in Table 5.

In light of the EFA findings, a bi-factor model was fitted in the CFA sample to compare a single general factor (i.e. all 10 items loading onto the same factor) with two distinct specific factors (e.g. a separate factor for the work items). Model fit indices demonstrated acceptable model fit (RMSEA = 0.06, CFI = 0.99 and SRMR = 0.02). The bi-factor model showed that the items loaded more highly onto the general factor than their specific factors (explained common variance [ECV] = 0.85), suggesting that the ILQI is unidimensional and a total score can be used. The bi-factor solution is presented in Fig. 3.

Reliability and Validity of Scores

When items of the ILQI were tested, Cronbach’s alpha was > 0.90, suggesting the items within the unidimensional score were interrelated and were working together to measure the same underlying construct [30] (Table 6).

Using a threshold of 0.90, the Spearman–Brown prophecy formula suggested that no more than three missing items are recommended for total score creation. Therefore, only calculating the ILQI total (alternative) score where a minimum of 7/10 items are completed is recommended.

Convergent validity was assessed between the ILQI total (alternative) score and individual

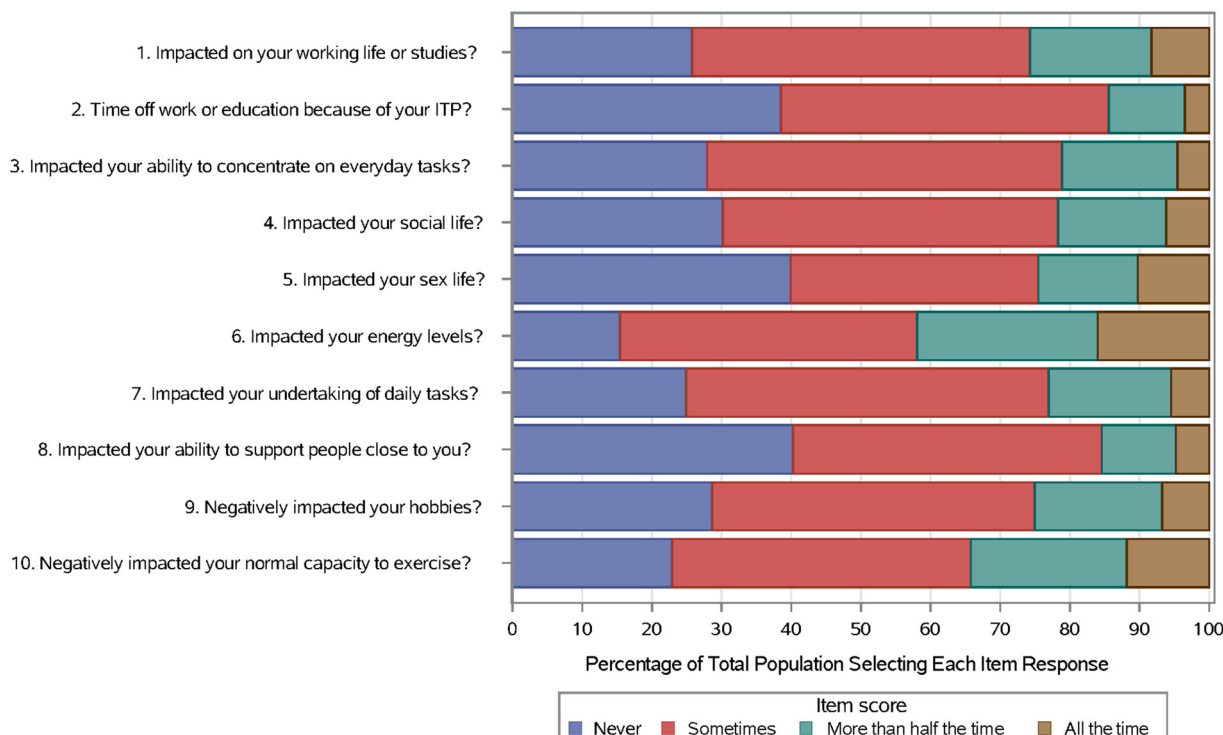


Fig. 2 Item response distributions

items and other clinical and demographic single questions, completed as part of I-WISH. As hypothesised, there were large correlations (≥ 0.50) [24] between the ILQI items 1 and 2 (e.g. assessing impact of ITP on work or study) with the demographic question assessing work productivity reduction ($r = 0.58$ – 0.68). Large correlations were also identified between the ILQI total (alternative) score and the emotional well-being question ($r = 0.65$). The daily activity reduction question correlated strongly with the ILQI total (alternative) score ($r = 0.72$) and also most items on the ILQI (items 3–4 and 6–10, $r > 0.50$). The question on current health was not as strongly correlated as hypothesised but was still within an acceptable (moderate) range ($r = 0.41$). ILQI item 5, assessing impact on sex life, continued to perform poorly (all r values < 0.30), potentially because of the number of patients who preferred not to answer this item ($n = 333$, 22.10%) and as a convergent measure specifically measuring sexual functioning was not included for correlation.

In known-groups analysis, ILQI total (alternative) scores were compared among distinct groups that were expected to have different ILQI scores (Table 7). All assessments showed that mean ILQI total (alternative) scores increased monotonically and differentiated between individuals who would be expected to have higher or lower HRQoL. A large between group effect size difference (> 0.80) was observed between most groups and all differences between consecutive groups were statistically significant ($p < 0.0001$).

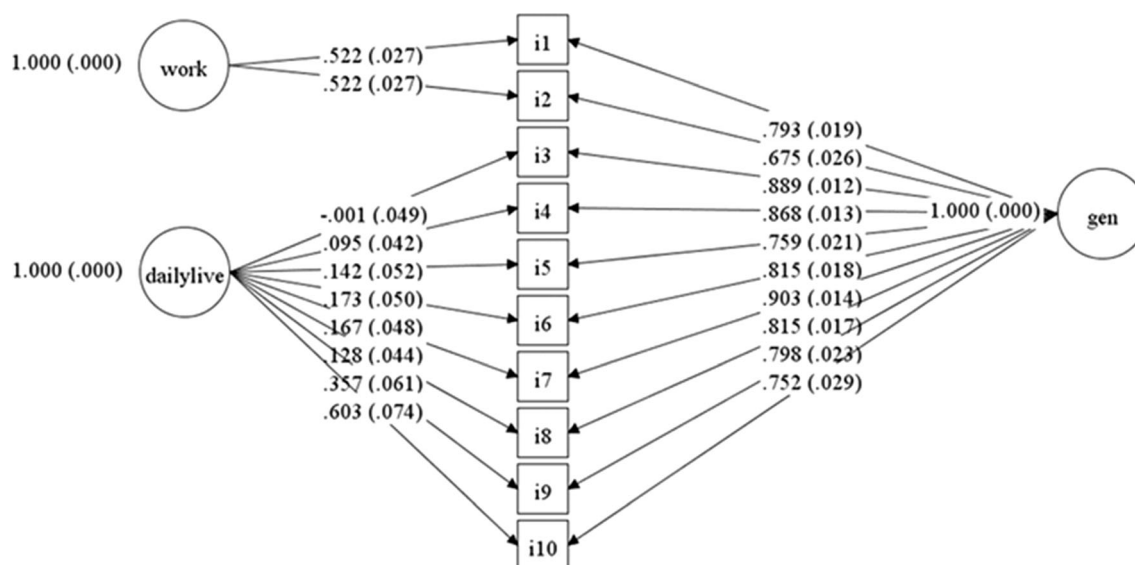
Potential thresholds were evaluated by finding an optional cut-point using ROC curves where sensitivity and specificity were both maximised. All ROC curves had an AUC significantly greater than 0.5. Optimal thresholds to discriminate between low and moderate severity on each anchor measure were current health (18, sensitivity = 0.596, specificity = 0.687); impact on emotional well-being (16, sensitivity = 0.721, specificity = 0.738); effect on daily activity reduction (19, sensitivity = 0.728, specificity = 0.829). Optimal thresholds to

Table 4 Inter-item correlations

Item	1	2	3	4	5	6	7	8	9	10
1. How often has your ITP impacted on your working life or studies?	1.00									
2. How often have you taken time off work or education because of your ITP?	0.8042	1.00								
3. How often has your ITP impacted your ability to concentrate on everyday tasks?	0.4947	0.4161	1.00							
4. How often has your ITP impacted your social life?	0.4577	0.3881	0.7742	1.00						
5. How often has your ITP impacted your sex life?	0.2205	0.2485	0.2854	0.3571	1.00					
6. How often has your ITP impacted your energy levels?	0.4504	0.3217	0.7009	0.6912	0.295	1.00				
7. How often has your ITP impacted your undertaking of daily tasks?	0.4718	0.3945	0.8023	0.7644	0.2881	0.8001	1.00			
8. How often has your ITP impacted your ability to support people close to you?	0.4355	0.3934	0.744	0.7094	0.2685	0.6536	0.7895	1.00		
9. How often has your ITP negatively impacted your hobbies?	0.4528	0.3707	0.7077	0.7491	0.3249	0.6694	0.7675	0.7249	1.00	
10. How often has your ITP negatively impacted your normal capacity to exercise?	0.4279	0.3625	0.6604	0.7009	0.3038	0.7229	0.7558	0.665	0.798	1.00

Table 5 Factor loadings for the EFA

Item	Unidimensional Factor 1	Two-factor	
		Factor 1	Factor 2
Factor correlations			0.779
1. How often has your ITP impacted on your working life or studies?	0.874	0.852	0.104
2. How often have you taken time off work or education because of your ITP?	0.849	0.932	− 0.004
3. How often has your ITP impacted your ability to concentrate on everyday tasks?	0.864	0.277	0.639
4. How often has your ITP impacted your social life?	0.842	0.071	0.792
5. How often has your ITP impacted your sex life?	0.783	0.011	0.787
6. How often has your ITP impacted your energy levels?	0.811	− 0.022	0.839
7. How often has your ITP impacted your undertaking of daily tasks?	0.899	0.04	0.878
8. How often has your ITP impacted your ability to support people close to you?	0.838	0.102	0.762
9. How often has your ITP negatively impacted your hobbies?	0.847	− 0.101	0.944
10. How often has your ITP negatively impacted your normal capacity to exercise?	0.820	− 0.079	0.897

**Fig. 3** CFA bi-factor solution

discriminate between moderate and high severity on each anchor measure were current health (25, sensitivity = 0.743, specificity = 0.547); impact on emotional well-being

(22, sensitivity = 0.703, specificity = 0.685); effect on daily activity reduction (25, sensitivity = 0.727, specificity = 0.576).

Table 6 Internal consistency reliability for the ILQI score

	Cronbach's alpha ILQI score (<i>N</i> = 1506)
Cronbach's alpha (total)	0.933
Cronbach's alpha if variable deleted	
1. How often has your ITP impacted on your working life or studies?	0.926
2. How often have you taken time off work or education because of your ITP?	0.930
3. How often has your ITP impacted your ability to concentrate on everyday tasks?	0.924
4. How often has your ITP impacted your social life?	0.925
5. How often has your ITP impacted your sex life?	0.929
6. How often has your ITP impacted your energy levels?	0.927
7. How often has your ITP impacted your undertaking of daily tasks?	0.923
8. How often has your ITP impacted your ability to support people close to you?	0.927
9. How often has your ITP negatively impacted your hobbies?	0.925
10. How often has your ITP negatively impacted your normal capacity to exercise?	0.926

DISCUSSION

Improving HRQoL and patient outcomes are key considerations when evaluating the impact of disease, particularly in ITP, where patients may still be experiencing significantly impaired HRQoL despite an adequate platelet count [6, 7, 12, 13, 31]. As such, discussions between patients and physicians are essential to ensure patients' HRQoL is considered and that ultimately a shared approach is taken towards treatment decision-making [31]. While the importance of a shared decision-making framework is heavily propagated in management of paediatric patients with ITP [32–34], there are fewer tools and resources to support it within care of adult patients with ITP. Notably, it is acknowledged although there are clinical outcome assessment tools available to assess symptoms of ITP in adults [31, 35] there is currently no disease-specific tool designed for use in clinical practice, with evidence of validity and reliability, which takes account of patient HRQoL to aid discussions between adult patients and physicians. The original ILQI, developed only by clinical experts, was refined, using methods in line with best practice

guidelines for the development and validation of PRO instruments [36–38]. Specifically, appropriate, in-depth qualitative work and rigorous psychometric methods were used to ensure the content validity of the ILQI.

The results presented here confirm that the ILQI is a valid and reliable PRO measure of HRQoL in patients with ITP. For six items, item response distributions were skewed towards the lower end of the scale (indicating high HRQoL); however, this is not a concern for the performance of the items, as the full range of response options were used for every item and other factors, such as disease severity, need to be considered. Despite the large ceiling effects for item 8, the concept which it assesses (support) emerged as important to patients with ITP in the qualitative research, suggesting that this item is more relevant to patients with more severe disease. Good inter-item correlations were identified between item-pairs that were expected to be highly correlated and no items were considered redundant. Considering the relevance of these concepts in the qualitative research, these items were retained for the remainder of the analyses in order to fully assess their fit. Consideration of the qualitative

Table 7 Known-groups analyses for the ILQI total (alternative) score

	Known-grouping	<i>n</i>	Mean (SD)	Between groups effect size	Linear trends <i>p</i> value	Between groups <i>p</i> value
Current health	Poor: 1–2	75	27.10 (7.957)		< 0.0001	
	Moderate: 3–5	962	21.78 (6.557)	0.80		< 0.0001
	Good: 6–7	465	17.40 (6.431)	1.46		< 0.0001
Impact on emotional well-being	Low impact: 1–2	301	14.32 (4.687)		< 0.0001	
	Moderate impact: 3–5	789	20.14 (5.638)	– 1.08		< 0.0001
	Large impact: 6–7	416	26.34 (6.425)	– 2.09		< 0.0001
Effect on daily activity reduction	Low impact: 0–3	736	16.13 (5.079)		< 0.0001	
	Moderate impact: 4–6	428	23.16 (4.731)	– 1.42		< 0.0001
	Large impact: 7–10	342	27.42 (6.069)	– 2.09		< 0.0001
Fatigue	No	754	18.44 (6.563)		N/A	
	Yes	752	22.95 (6.764)	– 0.68		< 0.0001
Depression	No	1236	19.64 (6.548)		N/A	
	Yes	270	25.52 (7.176)	– 0.88		< 0.0001

research and item content suggested that all the items were conceptually distinct.

EFA and CFA results supported a unidimensional solution, two-factor solution and a bi-factor solution with two distinct specific factors. However, the CFA bi-factor model showed that the items loaded more highly onto the general factor than the specific factor, suggesting that the ILQI is essentially unidimensional and the ILQI total (alternative) score can be used. Examining Cronbach's alpha with each item removed in turn found that the ILQI internal consistency was most reliable when all 10 items were included in the score. Results for convergent validity were strong between the ILQI and items assessing work productivity reduction, emotional well-being and daily activity reduction and there was a moderate correlation with current health. The item assessing impact on sex life continued to perform poorly, potentially because of the number of patients who did not answer this item and the lack of a convergent measure specifically assessing sexual

functioning. However, given the qualitative relevance of this item to patients, and the fact that it still contributed to the internal consistency reliability, it was retained. Known-groups analyses confirmed that ILQI total (alternative) scores were capable of distinguishing between clinically distinct groups.

The ROC-based analyses suggested that the existing thresholds were not representative and instead a threshold of around 17 points should be used to detect 'impaired HRQoL', and a threshold of around 23–25 points to detect 'significantly impaired HRQoL'.

Exploratory DIF analyses generated hypotheses regarding possible cultural differences in the way patients responded to specific items. Further linguistic and cultural validation is currently being conducted, which will explore the hypotheses arising from this analysis in greater detail through qualitative methods as recommended [15].

Study Limitations and Further Research

Several limitations should be considered when interpreting the results of this study. Firstly, patients were recruited via physicians and patient support groups as part of a larger survey/study. It must be acknowledged that patients enrolled via support groups may not be representative of the overall patient population seen in clinic. One concern is that patients who are engaged in support groups are typically those with more severe disease manifestation, and who are in need of additional support to manage their condition [39]. However, that was not the case in this sample as 85.2% of patients self-reported their current health state to be good or excellent. This relatively healthy sample should also be taken into consideration when interpreting these results.

Another potential limitation is that I-WISH did not include an existing validated measure which could be used as an established clinically significant distress variable. Correlations were assessed between the ILQI items/scores and clinical and demographic questions and there was no appropriate single-item assessment of sexual functioning.

The practicalities of administering a 10-item PRO measure in clinical practice and the time taken to discuss the findings with patients could be a challenge. As a result of time restrictions in clinic appointments, it may be more appropriate to identify a smaller set of questions to assess the impact of ITP on HRQoL. To address this, item response theory (IRT) analysis is currently being conducted to assess the discriminatory properties of the items and further assess the structure of the ILQI, with the aim of developing a more accurate scoring algorithm.

It is recommended that further validation work is conducted to assess the reliability of the ILQI longitudinally, to evaluate a change in score that would represent a minimal clinically important difference.

CONCLUSIONS

These findings provide evidence that the ILQI has acceptable psychometric properties as a PRO

instrument in patients with ITP. The cut-off scores derived from the work described in this paper help to optimally discriminate between severity groups and will aid patient-centred treatment decision-making between patients and physicians. The psychometric evaluation demonstrates that it is fit for measuring the impact of ITP on HRQoL in clinical practice. The adoption of the ILQI in routine care should improve consistency of patient-centred decision-making and may lead to better outcomes for those patients whose HRQoL has been negatively affected by their ITP. The intention is that the ILQI can be completed on paper or online, either before or during a consultation, to establish a patient's current HRQoL status and stimulate further discussion into the underlying causes of poor HRQoL, during the clinical interaction. The ongoing linguistic and cultural validation and IRT analysis aims to provide further evidence of validity and refine this measure.

ACKNOWLEDGEMENTS

We extend enormous thanks to the patients who took part in the I-WISH survey. We gratefully thank the ITP Support Association for their support. We also thank Nicola Bonner and Philip Griffiths, of Adelphi Values, who conducted the psychometric analyses, funded by Novartis Pharmaceuticals.

Funding. Adelphi Value were paid by Novartis Pharmaceuticals to conduct the research. Novartis funded the journal's rapid service fee.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Authors Contributions. Laura Grant performed research, analysed data and wrote the paper. Ricardo Viana and Denise D'Alessio

designed the research and wrote the paper. Nichola Cooper, Donald Arnold, Mervyn Morgan, Drew Provan, Adam Cuker, Quentin A Hill, Yoshiaki Tomiyama and Waleed Ghanima wrote the paper.

Disclosures. Laura Grant works for Adelphi Values who were paid by Novartis Pharmaceuticals to conduct the research. Ricardo Viana is an employee and stockholder of Novartis Pharma AG. Denise D'Alessio is an employee and stockholder of Novartis Pharmaceuticals. Nichola Cooper, Donald Arnold, Mervyn Morgan, Drew Provan, Adam Cuker, Quentin A Hill, Yoshiaki Tomiyama and Waleed Ghanima have nothing to disclose.

Compliance with Ethics Guidelines. The study was conducted in accordance with the Helsinki Declaration of 1964, and its later amendments. Survey materials and the study protocol were reviewed and approved by central institutional review boards (IRB) in both Europe and North America. All participants provided informed consent prior to completing the survey. No participant identifiable information is included in the manuscript.

Data Availability. The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Prior Presentation. Results were presented at American Society of Hematology conference 2019 and abstract published in *Blood* Volume 134, Issue Supplement 1, November 13 2019, title, "The Psychometric Properties of the ITP Life Quality Index Assessed in a Large Multinational "Real-World" Cohort of Immune Thrombocytopaenia Patients".

Open Access. This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and

indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. George JN, Mathias SD, Go RS, et al. Improved quality of life for romiplostim-treated patients with chronic immune thrombocytopenic purpura: results from two randomized, placebo-controlled trials. *Br J Haematol.* 2009;144(3):409–15.
2. Blickstein D. Treatment of immune thrombocytopenic purpura in adults: update. *Harefuah.* 2019;158(3):196–9.
3. Despotovic JM, Bussel JB. Immune thrombocytopenia (ITP). In: Michelson A, Cattaneo M, Frelinger A, Newman P, editors. *Platelets.* London: Elsevier; 2019. p. 707–24.
4. Mithoowani S, Arnold DM. First-line therapy for immune thrombocytopenia. *Hämostaseologie.* 2019;39(3):259–65.
5. Trotter P, Hill QA. Immune thrombocytopenia: improving quality of life and patient outcomes. *Patient Relat Outcome Meas.* 2018;9:369.
6. Cooper N, Kruse A, Kruse C, et al. Immune thrombocytopenia (ITP) World Impact Survey (iWISH): patient and physician perceptions of diagnosis, signs and symptoms, and treatment. *Am J Hematol.* 2021;96:188–98.
7. Cooper N, Kruse A, Kruse C, et al. Immune thrombocytopenia (ITP) World Impact Survey (iWISH): patient and physician perceptions of diagnosis, signs and symptoms, and treatment. *Am J Hematol.* 2021;96(2):188–98.
8. Sestøl HG, Trangbæk SM, Bussel JB, Frederiksen H. Health-related quality of life in adult primary immune thrombocytopenia. *Expert Rev Hematol.* 2018;11(12):975–85.

9. Provan D, Arnold DM, Bussel JB, et al. Updated international consensus report on the investigation and management of primary immune thrombocytopenia. *Blood Adv.* 2019;3(22):3780–817.
10. Griffiths P, Grant L, Bonner N, et al. The psychometric properties of the ITP Life Quality Index assessed in a large multinational “Real-World” cohort of immune thrombocytopenia patients. Washington, DC: American Society of Hematology; 2019.
11. Grant L, Bonner N, Harper A, et al., editors. Qualitative research to support the content validity of the Immune Thrombocytopenia (ITP) Life Quality Index (ILQI). *Br J Haematol.* 2018. Wiley 111 River St, Hoboken 07030-5774, NJ USA.
12. Cooper N, Kruse A, Kruse C, et al. Immune thrombocytopenia (ITP) World Impact Survey (I-WISH): impact of ITP on health-related quality of life. *Am J Hematol.* 2021;96(2):188–198.
13. Cooper N, Kruse A, Kruse C, et al. Immune thrombocytopenia (ITP) World Impact Survey (I-WISH): impact of ITP on health-related quality of life. *Am J Hematol.* 2021;96(2):199–207.
14. Reilly MC. Development of the Work Productivity and Activity Impairment (WPAI) Questionnaire. New York: Reilly Associates; 2008.
15. Scott NW, Fayers PM, Aaronson NK, et al. Differential item functioning (DIF) analyses of health-related quality of life instruments using logistic regression. *Health Qual Life Outcomes.* 2010;8(1):1–9.
16. Scott NW, Fayers PM, Aaronson NK, et al. A simulation study provided sample size guidance for differential item functioning (DIF) studies using short scales. *J Clin Epidemiol.* 2009;62(3):288–95.
17. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model Multidiscip J.* 1999;6(1):1–55.
18. Kline RB. Principles and practice of structural equation modeling. London: Guilford; 2015.
19. Shi D, Lee T, Maydeu-Olivares A. Understanding the model size effect on SEM fit indices. *Educ Psychol Meas.* 2019;79(2):310–34.
20. Nunnally JC. Psychometric theory 3E. New York: Tata McGraw-Hill Education; 1994.
21. Cohen J. The *t* test for means. *Stat Power Anal Behav Sci.* 1988;2:20–6.
22. Froud R, Abel G. Using ROC curves to choose minimally important change thresholds when sensitivity and specificity are valued equally: the forgotten lesson of Pythagoras. Theoretical considerations and an example application of change in health status. *PLoS ONE.* 2014;9(12):e114468.
23. Rice ME, Harris GT. Comparing effect sizes in follow-up studies: ROC area, Cohen’s *d*, and *r*. *Law Hum Behav.* 2005;29(5):615–20.
24. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale: Erlbaum; 1988.
25. Thissen DE, Wainer HE. Test scoring. Mahwah: Erlbaum; 2001.
26. Kruse C, Kruse A, Watson S, et al. Patients with immune thrombocytopenia (ITP) frequently experience severe fatigue but is it under-recognized by physicians: results from the ITP World Impact Survey (I-WISH). *Blood.* 2018;132(Supplement 1):2273.
27. Bussel J, Tomiyama Y, Michel M, et al. PF712 physicians’ perceptions on causes of primary and secondary ITP and leading causes of misdiagnosis: results from the ITP World Impact Survey (I-WISH). *HemaSphere.* 2019;3(S1):311.
28. Cooper N, Kruse C, Kruse A, et al. PF714 patient perceptions on splenectomy outcomes: results from the ITP World Impact Survey (I-WISH). *HemaSphere.* 2019;3(S1):312.
29. Cohen J. Set correlation and contingency tables. *Appl Psychol Meas.* 1988;12(4):425–34.
30. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34–42.
31. Terrell DR, Neunert CE, Cooper N, et al. Immune thrombocytopenia (ITP): current limitations in patient management. *Medicina (Kaunas).* 2020;56(12):667.
32. Beck CE, Boydell KM, Stasiulis E, et al. Shared decision making in the management of children with newly diagnosed immune thrombocytopenia. *J Pediatr Hematol Oncol.* 2014;36(7):559–65.
33. Grace RF, Despotovic JM, Bennett CM, et al. Physician decision making in selection of second-line treatments in immune thrombocytopenia in children. *Am J Hematol.* 2018;93(7):882–8.
34. Kochhar MK, Grilo S, Robbins L, Catalozzi M, Neunert C. Decision aid to support shared decision making in pediatric refractory immune thrombocytopenia. *Blood.* 2020;136(Supplement 1):47–8.

-
35. Akin S, Haznedaroglu IC. Clinical variations between three different causes of thrombocytopenia. *Niger J Clin Pract.* 2021;24(1):17–20.
 36. US Food and Drug Administration. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. *Fed Regist.* 2009;74(235):65132–3.
 37. Lasch KE, Marquis P, Vigneux M, et al. PRO development: rigorous qualitative research as the crucial foundation. *Qual Life Res.* 2010;19(8):1087–96.
 38. Patrick DL, Burke LB, Powers JH, et al. Patient-reported outcomes to support medical product labeling claims: FDA perspective. *Value Health.* 2007;10:S125–37.
 39. Owen JE, Boxley L, Goldstein MS, Lee JH, Breen N, Rowland JH. Use of health-related online support groups: population data from the California health interview survey complementary and alternative medicine study. *J Comput-Mediat Commun.* 2010;15(3):427–46.