














ORIGINAL ARTICLE

Psychometrics of the patient-reported outcomes measurement information system measures in hemophilia: the applicability of the pediatric item banks

Lorynn Teela^{1,2,3}  | Michiel A. J. Luijten^{1,3,4,5}  | Isolde A. R. Kuijlaars⁶   |
 Tessa C. M. van Gastel^{1,3,7} | Evelien S. van Hoorn⁸  | Samantha C. Gouw⁹  |
 Karin C. J. Fijnvandraat⁹  | Kathelijin Fischer⁶  | Marjon H. Cnossen¹⁰   |
 Sasja Andeweg¹⁰ | Carolien van der Velden – van 't Hoff¹⁰ | Corinne Liem¹¹ |
 Margreet E. Jansen-Zijlstra¹² | Marjolein Peters⁹  | Lotte Haverman^{1,2,3}  

¹Amsterdam UMC location University of Amsterdam, Emma Children's Hospital, Child and Adolescent Psychiatry & Psychosocial Care, Amsterdam, The Netherlands

²Amsterdam Public Health, Mental health and Digital health, Amsterdam, The Netherlands

³Amsterdam Reproduction and Development, Child development, Amsterdam, The Netherlands

⁴Amsterdam Public Health, Mental health and Methodology, Amsterdam, The Netherlands

⁵Epidemiology and Data Science, Amsterdam UMC location Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

⁶Centre for Benign Haematology, Thrombosis and Haemostasis, Van Creveldkliniek, University Medical Centre Utrecht, Utrecht University, Utrecht, The Netherlands

⁷Amsterdam Public Health, Mental health and Health Behaviours & Chronic Diseases, Amsterdam, The Netherlands

⁸Department of Public Health, Erasmus MC, Erasmus University Medical Centre Rotterdam, The Netherlands

⁹Amsterdam UMC location University of Amsterdam, Emma Children's Hospital, Paediatric Haematology, Amsterdam, The Netherlands

¹⁰Department of Paediatric Haematology, Erasmus MC Sophia Children's Hospital, Erasmus University Medical Centre Rotterdam, The Netherlands

¹¹Department of Pediatric Haemato-Oncology, Radboud University Medical Centre, Nijmegen, The Netherlands

¹²Department of Pediatric Haematology, University Medical Centre Groningen, Groningen, The Netherlands

Correspondence

Lotte Haverman, Child and Adolescent Psychiatry & Psychosocial Care, G8-136 Emma Children's Hospital/Amsterdam UMC, Postbox 22660, 1100 DD Amsterdam, The Netherlands.
 Email: l.haverman@amsterdamumc.nl

Handling Editor: Dr Bethany Samuelson Bannow

Abstract

Background: The use of patient-reported outcomes measures (PROMs) is important in hemophilia care, as it facilitates communication between patients and clinicians and promotes patient-centered care. Currently, a variety of PROMs with insufficient psychometric properties are used. Patient-reported outcomes measurement information system (PROMIS) measures, including Computer Adaptive Tests, were designed to measure generically and more efficiently and, therefore, are an alternative for the existing PROMs.

Objectives: To assess the feasibility, measurement properties, and outcomes of 8 PROMIS pediatric measures for boys with hemophilia.

Methods: In this multicenter study, boys with hemophilia completed 8 PROMIS measures and 2 legacy instruments. Feasibility was determined by the number of completed items and floor or ceiling effects (percentage of participants that achieved the lowest or highest possible score). Reliability was assessed as the percentage of scores with a

$SE \leq 4.5$. Construct validity was evaluated by comparing the PROMIS measures with the legacy instruments. Mean PROMIS T-scores were calculated and compared with the Dutch general population.

Results: In total, 77 boys with hemophilia participated. Reliability was good for almost all PROMIS measures and legacy instruments. The total number of completed items varied from 49 to 90 for the PROMIS pediatric measures, while the legacy instruments contained 117 to 130 items. Floor and ceiling effects were observed in both the PROMIS measures (0-39.5%) and legacy instruments (0-66.7%), but were higher for the legacy instruments.

Conclusions: The PROMIS pediatric measures are feasible to use for boys with hemophilia. With the use of the PROMIS measures in clinical care and research, a step toward worldwide standardization of PROM administration can be taken.

KEYWORDS

feasibility studies, hemophilia, patient-reported outcome measures, pediatrics, psychometrics

Essentials

- This study investigates the patient-reported outcomes measurement information system (PROMIS).
- This is a multicenter study in 76 Dutch children with hemophilia.
- The PROMIS pediatric measures are feasible to use for boys with hemophilia.
- The PROMIS pediatric measures are valid alternatives to the well-known legacy instruments.

1 | INTRODUCTION

Hemophilia A or B are x-linked bleeding disorders that are caused by a deficiency of the coagulation proteins, factor (F) VIII (hemophilia A) or FIX (hemophilia B), resulting in excessive bleeding typically in joints and muscles, spontaneously or after minor trauma. The risk of bleeding is related to the severity of the factor deficiency, and repeated bleeds can cause pain, functional impairment, and acute and long-term disabilities, especially when treated inadequately [1-3]. In recent decades, the treatment of hemophilia has greatly improved. In children with a severe phenotype of hemophilia, the treatment is now mainly focused on the prevention of bleeding by prophylactic therapy with factor concentrates (eg, prophylaxis) or non-factor alternatives (eg, emicizumab) [1,4-6].

With these treatment advancements, health outcomes and the health-related quality of life (HRQoL) of children with hemophilia have significantly improved. Children now have a near-normal life expectancy and HRQoL, experience a lower treatment burden, and are less limited in activities of daily living [5,7-10]. However, hemophilia treatment still has an impact on the lives of these children. Therefore, comprehensive care focusing on both physical and psychosocial outcomes is standard in high-income countries [4,11,12]. The use of patient-reported outcomes measures (PROMs) are of important value in comprehensive care to gain insight into the consequences of hemophilia treatment [13].

PROMs are self-reported questionnaires that measure patients' perspectives on their health, well-being, and the impact of disease and treatment on their life [14,15]. PROMs can be used both at a group level to study differences between populations or to measure the effect of treatment modalities in clinical research, or at an individual level to increase awareness for patients' problems and concerns, facilitate communication, and to guide clinical decision-making [16-18]. PROMs can be disease-specific (ie, applicable to patients with a specific disease) or generic (ie, applicable to everyone, regardless of disease) and are preferably standardized and validated [19]. For measuring outcomes in children, it is important to have PROMs available for different age ranges and parent proxy questionnaires.

In hemophilia research, a wide variety of PROMs are used which makes comparisons difficult due to differences in content, age ranges, and scoring methods [13,15,20-22]. Specifically for pediatric hemophilia care, a wide variety of disease-specific PROMs (eg, CHO-KLAT, Haemo-QoL) are used without established psychometric properties to justify the use of these disease-specific PROMs in daily clinical practice [20]. For these reasons, standardization of outcomes and PROM administration in hemophilia care and research is essential, as described by Van Hoorn et al. [11,15,20,23]. Several initiatives have recently worked on core outcome sets for patients with hemophilia [24-26], resulting in the patient-reported outcomes measurement information system (PROMIS) being selected as one of the included measurement tools [26].

PROMIS provides a set of person-centered, standardized instruments to measure a broad range of health domains (physical, mental, and social health) in children [27,28]. In contrast to legacy instruments that are based on the *Classical Test Theory* (CTT), PROMIS measures were developed according to the *Item Response Theory* (IRT) [29,30]. An important advantage of the use of IRT is the option of Computerized Adaptive Testing (CAT) [28,29]. With CAT, items are offered to patients based on their previous answers. Consequently, PROMIS measures are shorter, items are more tailored to the patients' situation, and the measurement is more reliable in comparison to existing PROMs [28,30]. Recent studies showed that, in (young) adult patients with hemophilia, PROMIS measures are effective, reliable, and valid with low floor- and ceiling effects [15,31–33]. However, it is unclear if PROMIS instruments are also suitable for children with hemophilia. Therefore, the aim of this study is to evaluate the feasibility, measurement properties, and outcomes of 8 relevant PROMIS pediatric measures for boys with hemophilia.

2 | METHODS

2.1 | Study population and procedure

All boys (8–17 years) treated for mild to severe hemophilia A or B in one of the hemophilia treatment centers in The Netherlands (Amsterdam University Medical Centers, Van Creveldkliniek, Erasmus University Medical Center, Radboud University Medical Center, or the University Medical Center Groningen) were eligible to participate in this multicenter study. Between June 2021 and December 2021, patients were invited to participate by email and received a personal link to the study website (<https://promis.hetklikt.nu/hemofilie/>) of the KLIK PROM portal [34]. Caregivers were asked to complete a socio-demographic questionnaire, and children were asked to complete 8 PROMIS instruments and 2 legacy instruments (Haemophilia Quality of Life Questionnaire for Children [Haemo-QoL] and Pediatric Hemophilia Activities List [PedHAL]). Children with insufficient knowledge of the Dutch language or children who were unable to complete the PROMs were excluded, as determined by the treating clinician.

The Medical Ethics Committees of the participating centers approved this study. All participants signed online informed consent.

2.2 | Measurements

2.2.1 | Patient characteristics

The sociodemographic questionnaire included questions about the caregivers (eg, country of birth, educational level, marital status), the child (eg, position in family, school, sports), and clinical characteristics/variables (eg, type and severity of hemophilia, treatment, bleeding episodes, comorbidities).

2.2.2 | PROMIS pediatric measures

Six PROMIS pediatric measures were assessed as CAT: V2.0 Pain Interference [35], V2.0 Fatigue [36], V2.0 Anxiety [37], V2.0 Depressive Symptoms [37], V2.0 Mobility [38], and V2.0 Peer Relationships [39]. For 2 domains, no CAT was available; therefore, we used the fixed scales: V2.0 Anger 9a scale [40] and V1.0 Global Health scale (7+2) [41]. All PROMIS pediatric measures use a 7-day recall period. Items are scored on a 5-point Likert scale ranging from 1 (“never”) to 5 (“almost always”), except for the domains Mobility (ranging from “not able to” to “with no trouble”) and Global Health (response categories differ for each item, eg, ranging from “excellent” to “poor”). The CAT automatically stopped when the SE (of the estimate was ≤ 3.2 (90% reliability) and/or a maximum of 12 items was administered. PROMIS total scores were calculated by transforming the item scores into a T-score ranging from 0 to 100. For all PROMIS pediatric measures, higher scores represent more of the construct (eg, more pain interference or better peer relationships). The scores of the total scales were calculated with use of the *PROMIS Assessment Center Scoring Service* (https://www.assessmentcenter.net/ac_scoringervice).

2.2.3 | Legacy instruments

The Haemo-QoL is a widely-used disease-specific instrument developed for the assessment of HRQoL of children with hemophilia [42]. The Haemo-QoL consists of different age versions. For this study, we used the Dutch versions for children aged 8 to 12 years (64 items) and adolescents aged 13 to 16 years (including children aged 17 years; 77 items). The Haemo-QoL measures 10 domains (*Physical Health, Feeling, Attitude, Family, Friends, Coping, Other People, Sport and School, Dealing, and Treatment*), and 2 additional domains for the adolescent version (*Future and Relationship*). Items are disease-specific and ask about complaints due to hemophilia (eg, the past 4 weeks I was sad due to my hemophilia). The Haemo-QoL uses a 4-week recall period and items are scored on a 5-point Likert scale ranging from “never” to “always.” Positively formulated items were inversely recoded and sum scores were calculated for each domain. Sum scores were transferred to transformed domain and total scores ranging from 0 to 100. Lower scores indicating better HRQoL.

The PedHAL is a validated disease-specific instrument that assesses the self-reported limitations in activities and participation for children (4–18 years) with hemophilia [43]. The PedHAL consists of 53 items, distributed over 7 domains (sitting/kneeling/standing, functions of the legs, functions of the arms, use of transportation, self-care, household tasks, and leisure activities and sports). The PedHAL uses a recall period of a month (eg, in the previous month, did you have any difficulty, due to hemophilia with walking short distances). Items are scored on a 6-point Likert scale ranging from “impossible” to “never a problem,” and a response option “not applicable.” Domain scores and a summary score were calculated and converted to normalized scores ranging from 0 to 100, where higher scores represent better

functioning. No scores were calculated if >50% of the items on a domain were scored as “not applicable.”

2.3 | Statistical analyses

The Statistical Package for Social Sciences version 26.0 was used for all statistical analyses. Descriptive analyses (means and percentages) were performed to characterize the patients.

2.3.1 | Reliability and feasibility

Reliability was assessed for the PROMIS instruments under IRT and for the legacy instruments under CTT. In IRT modeling, each response pattern results in a T-score and an associated reliability (SE of measurement). An SE of ≤ 4.5 corresponds to a reliability of 80%, which has been considered the minimum acceptable level of reliability for group comparisons with the PROMIS pediatric measures [40]. To assess the reliability of the PROMIS pediatric measures, the percentage of T-scores with an SE ≤ 4.5 was calculated. Internal consistency estimates (Cronbach α) were calculated to assess the reliability of the legacy instruments through CTT.

To assess the feasibility of the instruments for use in clinical practice the number of items (for CAT: mean, minimum, maximum) that patients completed were described. In addition, floor and ceiling effects for all instruments were calculated. Floor and ceiling effects were presented as the percentage of participants who achieved the lowest or highest possible score, respectively. A floor or ceiling effect was considered present if the commonly accepted threshold of 15% was exceeded [44,45]. Both the number of completed items as well as the floor and ceiling effects were compared between the PROMIS pediatric measures and the legacy instruments.

2.3.2 | Construct validity

To evaluate the convergent validity of the PROMIS pediatric measures, hypotheses regarding the correlations between the PROMIS pediatric measures and the legacy instrument were formulated by researchers of the project group (Table 1) and tested. Moderate correlations (Spearman’s rho, 0.40-0.69 [46]) were expected between PROMIS Pain Interference and Haemo-QoL Physical Health, PROMIS Depressive Symptoms and Haemo-QoL Feeling, PROMIS Mobility and PedHAL, and PROMIS Global Health and Haemo-QoL total score. Weak correlations (Spearman’s rho, 0.10-0.39 [46]) were expected between PROMIS Anxiety, PROMIS Anger and Haemo-QoL Feeling, and between PROMIS Peer Relationships and Haemo-QoL Other Persons. Although the constructs of these measures were closely related, the content differs due to the disease-specific vs generic approach.

Although no differences were expected between subgroups of boys with hemophilia based on previous literature [47,48], secondary

analysis were performed comparing the mean PROMIS T-scores of the subgroups severe vs non-severe (mild and moderate) hemophilia.

2.3.3 | Outcomes

To determine which PROMIS pediatric measures were relevant for patients with hemophilia, mean T-scores were calculated and compared with Dutch reference data [49–53] from the general male population (8-18 years) using independent t-tests. In addition, transformed/normalized total and scale scores of the legacy instruments were calculated.

2.4 | Synthesis of the results

Comparisons between the PROMIS pediatric measures and the legacy instruments are described regarding the number of completed items, floor and ceiling effects, and reliability.

3 | RESULTS

3.1 | Patient characteristics

A total of 77 boys with hemophilia participated (response rate: 47.5%). Of these, 70 participants (90.9%) completed all PROMs. The data of one participant was excluded, because this participant ticked the first answer for almost all questions in the PROMs ($N = 76$).

Patient characteristics are shown in Table 2. The mean age was 13.5 years (range, 8-17 years). The majority of the participants (86.8%) had hemophilia A. In addition, 40.8% of the participants had a severe form of hemophilia. These participants were treated with prophylaxis with factor concentrates (19 participants) or with emicizumab (12 participants). Participants with a moderate form of hemophilia (18.4%) received prophylaxis with factor concentrates (5 participants) or on-demand treatment (in case of a bleed; 9 participants). On-demand treatment was used for all participants with a mild form of hemophilia (35.5%).

3.2 | Reliability and feasibility

Table 3 and Table 4 show data on the number of completed items, reliability, and floor and ceiling effects of the PROMIS pediatric measures and legacy instruments, respectively. The reliability of the PROMIS pediatric measures was excellent (>90% of the scores were reliable) or good (>70% of the scores were reliable) for almost all measures, except for the CAT mobility (56.2% of the scores was reliable). The reliability of the legacy instruments was excellent with Cronbach α ranging from 0.92 to 0.99. The mean number of completed items per PROMIS pediatric measure varied from 8.8 items (range, 5-12) for the item bank *Peer Relationships* to 11.6 items (range, 8-12) for

TABLE 1 (Predefined) correlations between the patient-reported outcomes measurement information system (PROMIS) pediatric measures and the legacy instruments.

PROMIS pediatric measures	Legacy instruments	Version (y)	Predefined hypothesized correlations	Spearman's correlation	Confirmed
Pain Interference	Haemo-QoL Physical Health	8-12	≥ 0.40	0.49	Yes
		13-17	≥ 0.40	0.42	Yes
Anxiety	Haemo-QoL Feeling	8-12	≥ 0.10	0.60	Yes
		13-17	≥ 0.10	0.35	Yes
Depressive Symptoms	Haemo-QoL Feeling	8-12	≥ 0.40	0.63	Yes
		13-17	≥ 0.40	0.29	No
Mobility	PedHAL	8-17	≥ 0.40	0.41	Yes
Peer Relationships	Haemo-QoL Other People	8-12	≥ -0.10	-0.33	Yes
		13-17	≥ -0.10	-0.04	No
Anger	Haemo-QoL Feeling	8-17	≥ 0.10	0.44	Yes
		13-17	≥ 0.10	0.48	Yes
Global Health	Haemo-QoL total score	8-12	≥ -0.40	-0.51	Yes
		13-17	≥ -0.40	-0.20	No

Predefined correlations were either weak (> 0.10) or moderate (> 0.40) based on the content of the items and the domains assessed. Haemo-QoL, Haemophilia Quality of Life Questionnaire for Children; PedHAL, Paediatric Haemophilia Activities List; PROMIS, patient-reported outcomes measurement information system.

the item bank *Anxiety*. For the Haemo-QoL, the number of completed items varied from 6 items for the subscale *Haemo-QoL Other People* to 8 items for the subscale *Haemo-QoL Feeling*. The total number of items for the Haemo-QoL were 64 (8-12 years) and 77 items (13-17 years). The PedHAL consisted of 53 items. The selected set of PROMIS pediatric measures contained an average of 80.4 items (range, 49-90), while the selected legacy instruments contained 117 items for children aged 8-12 years and 130 items for children aged 13-17 years. This means a reduction of items by 31% for patients aged 8-12 years and a reduction of 38% for patients aged 13-17 years.

Floor and ceiling effects were present in both the PROMIS pediatric measures and the legacy instruments. For the PROMIS pediatric measures, floor effects were observed in 4 CATs: *Pain Interference*, *Fatigue*, *Anxiety*, and *Depressive Symptoms*. A ceiling effect was observed in the CAT *Mobility*. In case of floor or ceiling effects, participants had to complete the maximum of 12 items. For the legacy instruments, floor effects were observed for the *Haemo-QoL Physical Health*, *Feeling*, and *Other People*. Ceiling effects were observed for the *Haemo-QoL Physical Health*, *Feeling*, *Other People*, and *PedHAL (total score)*.

3.3 | Construct validity

The correlations between the PROMIS pediatric measures and the legacy instruments are shown in [Table 1](#). Of the 13 hypothesized correlations for convergent validity, 10 correlations were confirmed. The correlations between PROMIS Depressive Symptoms and Haemo-QoL Feeling 13 to 17 years (weak correlation), PROMIS Peer Relationships and Haemo-QoL Other People 13 to 17 years (negligible correlation), and PROMIS Global Health and Haemo-QoL total scores 13 to 17 years (weak correlation) did not meet the predefined correlations.

Secondary analysis showed that boys with severe hemophilia reported more fatigue (41.2 vs 38.0, $P = .04$, $d = 0.42$) compared with boys with non-severe hemophilia.

3.4 | Outcomes

[Figure](#) shows the mean PROMIS T-scores for boys with hemophilia and the Dutch reference group. In comparison with the Dutch

TABLE 2 Patient characteristics (N = 76).

Characteristics	N	Mean (SD)
Sociodemographic characteristics		
Age (y)	76	13.5 (2.8)
	N	%
Age groups		
8-12 y	31	40.8
13-17 y	45	59.2
Country of birth parents		
Both parents born in The Netherlands	62	81.6
At least 1 parent born in foreign country	10	13.1
Unknown	4	5.3
Hemophilia characteristics		
	N	%
Type of hemophilia		
Hemophilia A	66	86.8
Hemophilia B	6	7.9
Unknown	4	5.3
Severity of hemophilia		
Mild (5-50%)	27	35.5
Moderate (2-5%)	14	18.4
Severe (<1%)	31	40.8
Unknown	4	5.3
Type of treatment hemophilia		
Prophylaxis with factor concentrates	24	31.5
Prophylaxis with emicizumab	12	15.8
On demand - in case of a bleed	36	47.4
Unknown	4	5.3
Inhibitor		
Current	0	0
Historically	9	11.8
No inhibitor/unknown	67	88.2

reference data, boys with hemophilia reported more pain interference ($P < .001$, mean difference = 3.85, $d = 0.60$), and they scored worse on the domain *Mobility* ($P < .001$, mean difference = -6.33, $d = -1.02$). In contrast, boys with hemophilia scored better on the domain *Peer Relationships*. On the other domains, no differences were found between boys with hemophilia and the Dutch reference group.

On the legacy instruments, boys with hemophilia scored a mean transformed total score of 20.2/24.5 (range, 0.4-91.6) on the Haemo-QoL (8-12 years and 13-17 years, respectively). On the PedHAL, boys with hemophilia scored a mean normalized score of 96.5 (range, 40-100) (Table 4).

3.5 | Synthesis of the results

Table 5 presents a synthesis of the results for the PROMIS pediatric measures and the legacy instruments.

4 | DISCUSSION

This study evaluated the feasibility, measurement properties, and outcomes of 8 PROMIS pediatric measures in boys with hemophilia. Almost all PROMIS pediatric measures were considered feasible and reliable for use in clinical hemophilia care. The number of completed items in the selected set of PROMIS measures was lower than that of the legacy instruments, resulting in a lower burden of completing PROMs. However, at domain level, the number of completed items was higher for the PROMIS pediatric measures, except for the measures *Mobility* and *Global Health*. Floor and ceiling effects of the PROMIS pediatric measures were substantially less than that of the legacy instruments. This implies that PROMIS measures adequately cover the range of functioning of boys with hemophilia. The reliability of the PROMIS pediatric measures and the legacy instruments was good, with exception of the PROMIS CAT *Mobility*.

4.1 | Validity

For testing convergent validity, we choose the widely-used disease-specific PROMs within hemophilia pediatric care and research (Haemo-QoL and PedHAL) [20]. These PROMs aim to measure the effect of hemophilia on daily life, and specifically ask if children experience symptoms like pain, sadness, or problems with friends due to their hemophilia (eg, I was angry because of my hemophilia). This is different from the PROMIS pediatric instruments that measure a generic domain of health and assume that symptoms can occur due to multifaceted reasons (eg, I was angry) [19]. Due to these different approaches, strong correlations were not expected, and it was hard to accurately assess convergent validity. For example, the *PROMIS peer relationships* item bank was expected to correlate minimally with the *Haemo-QoL Other People* scale as they assess different domains of social health. The *Other People* scale of Haemo-QoL relates more to the ability to participate in social roles due to hemophilia, whereas the *Peer Relationships* item bank relates to the overall quantity and quality of relationships with peers. Similarly, the *Haemo-QoL Feeling* scale does not cover the same unidimensional domains as measured by the PROMIS item banks. Nonetheless, most convergent validity hypotheses were met in both age groups, except for *PROMIS Peer Relationships* and *PROMIS Depressive Symptoms* item banks and the *Global Health* scale for children aged 13 to 17 years. Previous studies have shown that the subjective questioning of the *Global Health* scale ("How would you rate your own health?") may be influenced by social norms, which could be a possible explanation for a low correlation

TABLE 3 Number of completed items, reliability, floor and ceiling effects, and mean scores of the patient-reported outcomes measurement information system (PROMIS) pediatric measures.

PROMIS pediatric measures	No. of items			Reliability %	Floor %	Ceiling %	Mean T-score (σ)	N
	Mean	Minimum	Maximum					
Computerized Adaptive Tests								
Pain interference	9.2	3	12	100	39.5	0	42.1 (6.5)	76
Fatigue	11.4	8	12	100	16.2	0	39.9 (7.9)	74
Anxiety	11.6	8	12	100	28.0	0	42.3 (5.9)	75
Depressive Symptoms	10.3	4	12	100	22.7	0	43.9 (7.4)	75
Mobility	11.1	3	12	55.4	0	32.4	52.4 (6.2)	74
Peer Relationships	8.8	5	12	100	0	13.7	49.6 (7.5)	73
Scale								
Anger	9	9	9	100	9.5	0	43.6 (7.3)	74
Global Health	9	9	9	87.5	0	1.4	50.1 (7.9)	72

Reliability: scores were considered reliable as $SE \leq 4.5$. PROMIS, patient-reported outcomes measurement information system.

with the more objective questioning of the Haemo-QoL (total score of all subdomains), which relates much more to reported symptoms [53].

In addition, the correlation between the PROMIS pediatric measures and the legacy instruments could be negatively affected by the high floor and ceiling effects and the differences in recall period [33]. The PROMIS instruments use a recall period of 7 days, while the legacy instruments apply recall periods of 4 weeks/month [54].

A limitation of this study is that we were unable to directly compare the reliability of the PROMIS instruments and legacy instruments, due to the use of different measurement theories (IRT vs CTT). Results showed that both the PROMIS pediatric measures as

well as the legacy instruments measure reliably. However, higher floor and ceiling effects were found for the legacy instruments than for the PROMIS item banks (except for the *PROMIS Mobility* item bank) negatively affecting content validity and reliability. This is in accordance with previous studies on the PedHAL and Haemo-QoL instruments, where floor and ceiling effects were also found [42,43]. High floor and ceiling effects implicate that distinctive items are missing at the ends of the scale, making it difficult to distinguish patients with few or no complaints from each other [44], which results in an unreliable measurement for these patients. This also may explain the low reliability for the *PROMIS Mobility* item bank.

TABLE 4 Number of completed items, reliability, floor and ceiling effects, and mean scores of the legacy instruments.

Legacy instruments	Version (y)	No. of items	Reliability α	Floor %	Ceiling %	M (σ)	N
Haemo-QoL Physical Health							
	8-12	7	0.92	30.0	0	16.7 (23.2)	30
	13-17	7	0.95	43.2	5.4	15.5 (25.9)	37
Haemo-QoL Feeling							
	8-12	7	0.99	60.0	3.3	10.4 (25.7)	30
	13-17	8	0.97	51.4	5.4	12.9 (26.5)	37
Haemo-QoL Other People							
	8-12	6	0.97	66.7	3.3	9.3 (24.6)	30
	13-17	6	0.97	56.8	5.4	13.2 (26.7)	37
Haemo-QoL total score							
	8-12	64	0.95	0	0	20.2 (15.4)	30
	13-17	77	0.98	0	0	24.5 (20.2)	37
PedHAL (total score)							
	8-17	53	0.98	0	44.8	96.5 (9.4)	67

Data of 4 patients (Haemo-QoL, 13-17 years) were excluded as these patients experienced technical difficulties during completion. Haemo-QoL, Haemophilia Quality of Life Questionnaire for Children; PedHAL, Paediatric Haemophilia Activities List.

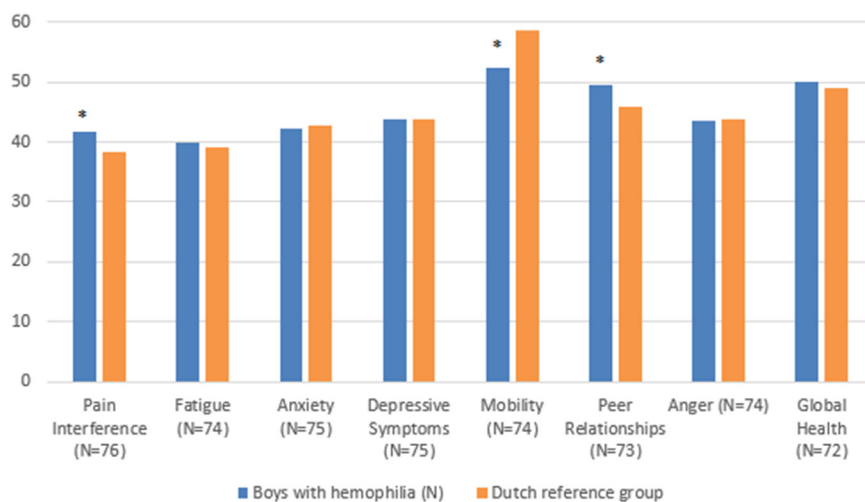


FIGURE Mean patient-reported outcomes measurement information system (PROMIS) T-scores for boys with hemophilia and the Dutch reference group.

4.2 | Health-related quality of life

The results of this study showed that the HRQoL of boys with hemophilia is comparable to the Dutch general population, except for the domains *Pain Interference* and *Mobility*. The high HRQoL found in this study is comparable to other studies assessing the HRQoL of boys with hemophilia with the legacy instruments [9,10,42]. Boys with a severe phenotype of hemophilia in The Netherlands experience few

joint bleeds because the annual bleeding rate is low due to adequate prophylactic therapy. It is therefore recommended to repeat this study in a group of boys with hemophilia in low-income countries with less access to effective treatment.

A limitation of this study is that as a measure of sociocultural determinants of the population, we did not have information on the race or ethnicity of participants, but did present information on place of birth of parents as a proxy for this.

TABLE 5 Comparison between the measurement properties of the patient-reported outcomes measurement information system (PROMIS) pediatric measures and legacy instruments.

PROMIS pediatric measures	Feasibility			Measurement properties	
	<i>N_items</i>	Floor ^a	Ceiling ^a	Reliability ^b	Convergent validity ^c
Pain Interference	9.2	-	+	+	+
Fatigue	11.4	+/-	+	+	n/a
Anxiety	11.6	+/-	+	+	+
Depressive Symptoms	10.3	+/-	+	+	+/-
Mobility	11.1	+	-	-	+
Peer Relationships	8.8	+	+	+	+/-
Anger	9	+	+	+	+
Global health	9	+	+	+	+/-
Legacy instruments					
Haemo-QoL Physical Health	7	-	+	+	
Haemo-QoL Feeling	7/8 ^d	-	+	+	
Haemo-QoL Other People	6	-	+	+	
Haemo-QoL total score	64/77 ^d	+	+	+	
PedHal (total score)	53	+	-	+	

Haemo-QoL, Haemophilia Quality of Life Questionnaire for Children; PedHAL, Paediatric Haemophilia Activities List; PROMIS, patient-reported outcomes measurement information system.

^aFloor/ceiling effect: + = <15% , +/- = 15-30%, - = ≥30%.

^bReliability: + = SE ≤ 4.5, - = SE > 4.5.

^cConvergent validity: + = predefined correlations are met, +/- = predefined correlations are partially met.

^dDifferent number of items for the age version 8 to 12 years and 13 to 17 years.

4.3 | Future research

The number of the PROMIS CAT items administered was still relatively high. The reason for this is that available items on the high or low end of the scale are limited and more difficult to measure reliably. Consequently, patients with no problems or complaints have to answer the maximum amount of items to reach the CAT stopping rule ($SE \leq 3.2$ and/or a maximum of 12 items). To reduce the burden of administration of PROMs for patients, initiatives are currently exploring the possibility to optimize these CAT stopping rules [55]. There also have been initiatives to shorten the legacy instruments [56,57].

5 | CONCLUSION

The PROMIS pediatric measures are reliable and feasible to use in hemophilia clinical care and research. Although, more research is needed to further reduce the burden of completing PROMs and to get more insight into the minimal important changes in patients with hemophilia. Innovative therapies are currently implemented or researched in clinical trials [4,5]. The need for reliable and valid instruments is crucial to measure the impact and cross-benefit of these innovative treatments. We conclude that the PROMIS measures are valid alternatives to the well-known legacy instruments, and importantly demonstrate lower floor and ceiling effects. With the use of generic PROMIS pediatric measures as used in our study, a leap can be made toward worldwide standardization of PROM administration, realizing comparisons between patient populations, the general population, patients from other disease groups, and other health care settings [23].

ACKNOWLEDGMENTS

We would like to thank all participants in this study. In addition, we acknowledge professor, Dr Caroline Terwee for her help in interpreting the results.

AUTHOR CONTRIBUTIONS

M.P. and L.H. conceived the study. L.T., I.A.R.K., E.Sv.H., S.C.G., K.C.J.F., K.F., and M.H.C. contributed to study design. Data collection was led by L.T. and supported by M.H.C., S.A., Cvd.Vvt.H., C.L., M.E.J.Z. L.T. performed the statistical analyses and M.A.J.L., T.C.Mv.G., M.P., LH contributed to the interpretation of the data and modification of statistical analyses. The first draft of the paper was written by L.T. M.P. and L.H. were responsible for the supervision. All authors critically revised the manuscript for intellectual content and approved the final version of the manuscript.

FUNDING

This study was funded with an unrestricted research grant from Pfizer.

RELATIONSHIP DISCLOSURE

MAJL and LH are members of the Dutch-Flemish PROMIS group. SCG received an unrestricted research grant from Sobi. The other authors have no conflict of interest regarding this work/project.

ETHICS STATEMENT

The study was approved by the Medical Ethic Research board from all participating centers. All procedures performed in this study were in accordance with the ethical standards of the international and/or national research committee (Medical Ethics Committee of the Amsterdam UMC - W19_349 # 21.111) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

INFORMED PATIENT CONSENT

Informed consent was provided by all patients for the use of their data for this study.

ORCID

Lorynn Teela  <https://orcid.org/0000-0002-8067-2792>

Michiel A. J. Luijten  <https://orcid.org/0000-0001-8016-2859>

Isolde A. R. Kuijlaars  <https://orcid.org/0000-0003-2920-2258>

Evelien S. van Hoorn  <https://orcid.org/0000-0001-9661-2746>

Samantha C. Gouw  <https://orcid.org/0000-0002-1957-4122>

Karin C. J. Fijnvandraat  <https://orcid.org/0000-0003-0904-4360>

Kathelijn Fischer  <https://orcid.org/0000-0001-7126-6613>

Marjon H. Cnossen  <https://orcid.org/0000-0003-1557-2995>

Marjolein Peters  <https://orcid.org/0000-0001-8419-7944>

Lotte Haverman  <https://orcid.org/0000-0001-7849-0562>

TWITTER

Isolde A.R. Kuijlaars  @IsoldeKuijlaars

Marjon H. Cnossen  @CnossenM

Lotte Haverman  @LotteHaverman

REFERENCES

- [1] Peyvandi F, Garagiola I, Young G. The past and future of haemophilia: diagnosis, treatments, and its complications. *Lancet*. 2016;388:187–97.
- [2] Boehlen F, Graf L, Berntorp E. Outcome measures in haemophilia: a systematic review. [Review]. *Eur J Haematol Suppl*. 2014;76:2–15.
- [3] Van Vulpen LFD, Holstein K, Martinoli C. Joint disease in haemophilia: pathophysiology, pain and imaging. *Haemophilia*. 2018;24(Suppl 6):44–9.
- [4] Srivastava A, Santagostino E, Dougall A, Kitchen S, Sutherland M, Pipe SW, et al. WFH guidelines for the management of hemophilia. *Haemophilia*. 2020;26(Suppl 6):1–158.
- [5] Balkaransingh P, Young G. Novel therapies and current clinical progress in hemophilia A. *Ther Adv Hematol*. 2018;9:49–61.
- [6] Oldenburg J, Mahlangu JN, Kim B, Schmitt C, Callaghan MU, Young G, et al. Efficacy of emicizumab prophylaxis in hemophilia A with inhibitors. *N Engl J Med*. 2017;377:809–18.
- [7] Shapiro S, Makris M. Haemophilia and ageing. *Br J Haematol*. 2019;184:712–20.
- [8] Versloot O, Timmer MA, de Kleijn P, Schuurings M, van Koppenhagen CF, van der Net J, et al. Sports participation and sports injuries in Dutch boys with haemophilia. *Scand J Med Sci Sports*. 2020;30:1256–64.

- [9] Bullinger M, von Mackensen S. Quality of life in children and families with bleeding disorders. *J Pediatr Hematol Oncol*. 2003;25(Suppl 1):64–7.
- [10] Kuijlaars IAR, van der Net J, Schutgens REG, Fischer K. The Paediatric Haemophilia Activities List (pedHAL) in routine assessment: changes over time, child-parent agreement and informative domains. *Haemophilia*. 2019;25:953–9.
- [11] Cassis FRMY, Querol F, Iorio A, Forsyth A. Psychosocial aspects of haemophilia: a systematic review of methodologies and findings. *Haemophilia*. 2012;18:101–14.
- [12] Hughes T, Brok-Kristensen M, Gargeya Y, Worsøe Lottrup AM, Bo Larsen A, Torres-Ortuño A, et al. "What more can we ask for?": an ethnographic study of challenges and possibilities for people living with haemophilia. *The J Haemophil Pract*. 2020;7:25–36.
- [13] Manco-Johnson MJ, Warren BB, Buckner TW, Funk SM, Wang M. Outcome measures in haemophilia: beyond ABR (annualized bleeding rate). *Haemophilia*. 2021;27(Suppl 3):87–95.
- [14] Weldring T, Smith SMS. Article commentary: patient-reported outcomes (pros) and patient-reported outcome measures (PROMs). *Health Serv Insights*. 2013;6.
- [15] Heesterbeek MR, Luijten MAJ, Gouw SC, Limperg PF, Fijnvandraat K, Coppens M, et al. Measuring anxiety and depression in young adult men with haemophilia using PROMIS. *Haemophilia*. 2022;28:e79–82.
- [16] Valderas JM, Kotzeva A, Espallargues M, Guyatt G, Ferrans CE, Halyard MY, et al. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Qual Life Res*. 2008;17:179–93.
- [17] Ishaque S, Karnon J, Chen G, Nair R, Salter AB. A systematic review of randomised controlled trials evaluating the use of patient-reported outcome measures (PROMs). *Qual Life Res*. 2019;28:567–92.
- [18] Bele S, Chugh A, Mohamed B, Teela L, Haverman L, Santana MJ. Patient-reported outcome measures in routine pediatric clinical care: a systematic review. *Front Pediatr*. 2020;8:364.
- [19] Terwee CB, Zuidgeest M, Vonkeman HE, Cella D, Haverman L, Roorda LD. Common patient-reported outcomes across ICHOM Standard Sets: the potential contribution of PROMIS(R). *BMC Med Inform Decis Mak*. 2021;21:259.
- [20] Limperg PF, Terwee CB, Young NL, Price VE, Gouw SC, Peters M, et al. Health-related quality of life questionnaires in individuals with haemophilia: a systematic review of their measurement properties. *Haemophilia*. 2017;23:497–510.
- [21] Timmer MA, Gouw SC, Feldman BM, Zwagemaker A, de Kleijn P, Pisters MF, et al. Measuring activities and participation in persons with haemophilia: a systematic review of commonly used instruments. *Haemophilia*. 2018;24:e33–49.
- [22] Gouw SC, Timmer MA, Srivastava A, de Kleijn P, Hilliard P, Peters M, et al. Measurement of joint health in persons with haemophilia: a systematic review of the measurement properties of haemophilia-specific instruments. *Haemophilia*. 2019;25:e1–10.
- [23] van Hoorn ES, Teela L, Kuijlaars IAR, Fischer K, Gouw SC, Cnossen MH, et al. Harmonizing patient-reported outcome measurements in inherited bleeding disorders with PROMIS. *Haemophilia*. 2023;29:357–61.
- [24] Iorio A, Skinner MW, Clearfield E, Messner D, Pierce GF, Witkop M, et al. Core outcome set for gene therapy in haemophilia: results of the core HEM multistakeholder project. *Haemophilia*. 2018;24:e167–72.
- [25] Dover S, Blanchette VS, Srivastava A, Fischer K, Abad A, Feldman BM. Clinical outcomes in hemophilia: towards development of a core set of standardized outcome measures for research. *Res Pract Thromb Haemost*. 2020;4:652–8.
- [26] van Balen EC, O'Mahony B, Cnossen MH, Dolan G, Blanchette VS, Fischer K, et al. Patient-relevant health outcomes for hemophilia care: development of an international standard outcomes set. *Res Pract Thromb Haemost*. 2021. <https://doi.org/10.1002/rth2.12488>
- [27] Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol*. 2010;63:1179–94.
- [28] Broderick JE, DeWitt EM, Rothrock N, Crane PK, Forrest CB. Advances in patient-reported outcomes: the NIH PROMIS® measures. *eGEMs (Wash DC)*. 2013;1:1015.
- [29] Cella D, Gershon R, Lai JS, Choi S. The future of outcomes measurement: Item banking, tailored short-forms, and computerized adaptive assessment. *Qual Life Res*. 2007;16(Suppl 1):133–41.
- [30] Fries JF, Witter J, Rose M, Cella D, Khanna D, Morgan-DeWitt E. Item response theory, computerized adaptive testing, and PROMIS: assessment of physical function. *J Rheumatol*. 2014;41:153–8.
- [31] Kuijlaars IAR, Teela L, van Vulpen LFD, Timmer MA, Coppens M, Gouw SC, et al. Generic PROMIS item banks in adults with hemophilia for patient-reported outcome assessment: feasibility, measurement properties, and relevance. *Res Pract Thromb Haemost*. 2021;5:e12621.
- [32] van Balen EC, Haverman L, Hassan S, Taal EM, Smit C, Driessens MH, et al. Validation of PROMIS Profile-29 in adults with hemophilia in the Netherlands. *J Thromb Haemost*. 2021;19:2687–701.
- [33] Barry V, Buckner TW, Lynch ME, Figueroa J, Mattis S, Stout ME, et al. An evaluation of PROMIS health domains in adults with haemophilia: a cross-sectional study. *Haemophilia*. 2021;27:375–82.
- [34] Haverman L, Engelen V, Van Rossum MA, Heymans HS, Grootenhuis MA. Monitoring health-related quality of life in paediatric practice: development of an innovative web-based application. *BMC Pediatr*. 2011;11:3.
- [35] Varni JW, Stucky BD, Thissen D, DeWitt EM, Irwin DE, Lai JS, et al. PROMIS pediatric pain interference scale: an item response theory analysis of the pediatric pain item bank. *J Pain*. 2010;11:1109–19.
- [36] Lai JS, Stucky BD, Thissen D, Varni JW, DeWitt EM, Irwin DE, et al. Development and psychometric properties of the PROMIS® pediatric fatigue item banks. *Qual Life Res*. 2013;22:2417–27.
- [37] Irwin DE, Stucky B, Langer MM, Thissen D, DeWitt EM, Lai JS, et al. An item response analysis of the pediatric PROMIS anxiety and depressive symptoms scales. *Qual Life Res*. 2010;19:595–607.
- [38] DeWitt EM, Stucky BD, Thissen D, Irwin DE, Langer M, Varni JW, et al. Construction of the eight-item patient-reported outcomes measurement information system pediatric physical function scales: built using item response theory. *J Clin Epidemiol*. 2011;64:794–804.
- [39] DeWalt DA, Thissen D, Stucky BD, Langer MM, Morgan Dewitt EM, Irwin DE, et al. PROMIS pediatric peer relationships scale: development of a peer relationships item bank as part of social health measurement. *Health Psychol*. 2013;32:1093–103.
- [40] Irwin DE, Stucky BD, Langer MM, Thissen D, DeWitt EM, Lai JS, et al. PROMIS pediatric anger scale: an item response theory analysis. *Qual Life Res*. 2012;21:697–706.
- [41] Forrest CB, Bevans KB, Pratiwadi R, Moon J, Teneralli RE, Minton JM, Tucker CA. Development of the PROMIS® Pediatric Global Health (PGH-7) Measure. *Qual Life Res*. 2014;23:1221–31.
- [42] von Mackensen S, Bullinger M, Haemo-QoL Group. Development and testing of an instrument to assess the Quality of Life of Children with Haemophilia in Europe (Haemo-QoL). *Haemophilia*. 2004;10(Suppl 1):17–25.
- [43] Groen WG, Van Der Net J, Helders PJM, Fischer K. Development and preliminary testing of a Paediatric Version of the Haemophilia Activities List (pedhal). *Haemophilia*. 2010;16:281–9.
- [44] Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement

- properties of health status questionnaires. *J Clin Epidemiol*. 2007;60:34–42.
- [45] Lim CR, Harris K, Dawson J, Beard DJ, Fitzpatrick R, Price AJ. Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. *BMJ Open*. 2015;5:e007765.
- [46] Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. *Anesth Analg*. 2018;126:1763–8.
- [47] Limperg PF, Joosten MMH, Fijnvandraat K, Peters M, Grootenhuys MA, Haverman L. Male gender, school attendance and sports participation are positively associated with health-related quality of life in children and adolescents with congenital bleeding disorders. *Haemophilia*. 2018;24:395–404.
- [48] McCusker PJ, Fischer K, Holzhauer S, Meunier S, Altisent C, Grainger JD, et al. International cross-cultural validation study of the Canadian haemophilia outcomes: kids' life assessment tool. *Haemophilia*. 2015;21:351–7.
- [49] Luijten MA, van Litsenburg RR, Terwee CB, Grootenhuys MA, Haverman L. Psychometric properties of the Patient-Reported Outcomes Measurement Information System (PROMIS®) pediatric item bank peer relationships in the Dutch general population. *Qual Life Res*. 2021;1–10.
- [50] Klaufus LH, Luijten MAJ, Verlinden E, van der Wal MF, Haverman L, Cuijpers P, et al. Psychometric properties of the Dutch-Flemish PROMIS® pediatric item banks Anxiety and Depressive Symptoms in a general population. *Qual Life Res*. 2021;30:2683–95.
- [51] Peersmann SHM, Luijten MAJ, Haverman L, Terwee CB, Grootenhuys MA, van Litsenburg RRL. Psychometric properties and CAT performance of the PROMIS pediatric sleep disturbance, sleep-related impairment, and fatigue item banks in Dutch children and adolescents. *Psychol Assess*. 2022;34:860–9.
- [52] van Muilekom MM, Luijten MAJ, van Litsenburg RRL, Grootenhuys MA, Terwee CB, Haverman L, et al. Psychometric properties of the Patient-Reported Outcomes Measurement Information System (PROMIS®) pediatric Anger scale in the general Dutch population. *Psychol Assess*. 2021;33:1261–6.
- [53] Luijten MAJ, Haverman L, van Litsenburg RRL, Roorda LD, Grootenhuys MA, Terwee CB. Advances in measuring pediatric overall health: the PROMIS® Pediatric Global Health scale (PGH-7). *Eur J Pediatr*. 2022;181:2117–25.
- [54] Coombes L, Bristowe K, Ellis-Smith C, Aworinde J, Fraser LK, Downing J, et al. Enhancing validity, reliability and participation in self-reported health outcome measurement for children and young people: a systematic review of recall period, response scale format, and administration modality. *Qual Life Res*. 2021;30:1803–32.
- [55] Kallen MA, Cook KF, Amtmann D, Knowlton E, Gershon RC. Grooming a CAT: customizing CAT administration rules to increase response efficiency in specific research and clinical settings. *Qual Life Res*. 2018;27:2403–13.
- [56] Pollak E, Mühlan H, VON Mackensen S, Bullinger M, HAEMO-QOL GROUP. The Haemo-QoL Index: developing a short measure for health-related quality of life assessment in children and adolescents with haemophilia. *Haemophilia*. 2006;12:384–92.
- [57] Kuijlaars IAR, van der Net J, Bouskill V, Hilliard P, Juodyte A, Khair K, et al. Shortening the paediatric Haemophilia Activities List (pedHAL) based on pooled data from international studies. *Haemophilia*. 2021;27:305–13.