Received: 23 July 2020 | Revised: 15 September 2020 | Accepted: 24 September 2020

DOI: 10.1111/hae.14180

Haemophilia WILEY

ORIGINAL ARTICLE

Muskuloskeletal

Evaluating international Haemophilia Joint Health Score (HJHS) results combined with expert opinion: Options for a shorter HJHS

Isolde A.R. Kuijlaars[1] | Janjaap van der Net[2] | Brian M. Feldman[3,4,5] | Magnus Aspdahl[6] | Melanie Bladen[7] | Wypke de Boer[8] | Rubén Cuesta-Barriuso[9,10,11] | Ruth E.D. Matlary[12] | Sharon M. Funk[13] | Pamela Hilliard[3] | Judy A. John[14] | Christine L. Kempton[15] | Piet de Kleijn[1] | Marilyn Manco-Johnson[13] | Pia Petrini[6] | Pradeep Poonnoose[16] | Jean St-Louis[17] | Sylvia Thomas[18] | Merel A. Timmer[1] | Sonata Saulyte Trakymiene[19] | Leo van Vlimmeren[20] | Kathelijn Fischer[1]

[1]Van Creveldkliniek, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

[2]Center for Child Development, Exercise and Physical Literacy, Children's Hospital of the University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

[3]Child Health Evaluative Sciences Program, Research Institute, The Hospital for Sick Children, Toronto, Ontario, Canada

[4]Department of Pediatrics, Faculty of Medicine and the Institute of Health Policy Management & Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada

[5]Division of Rheumatology, The Hospital for Sick Children, Toronto, Ontario, Canada

[6]Department of Pediatrics, Clinic of Coagulation Disorders, Karolinska University Hospital, Stockholm, Sweden

[7]Haemophilia Center, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK

[8]Department of Rehabilitation, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands

[9]Department of Physiotherapy, European University of Madrid, Madrid, Spain

[10]Royal Victoria Eugenia Foundation, Madrid, Spain

[11]Fishemo CEE, Spanish Federation of Hemophilia, Madrid, Spain

Abstract

Introduction: The Hemophilia Joint Health Score (HJHS) was developed to detect early changes in joint health in children and adolescents with haemophilia. The HJHS is considered by some to be too time consuming for clinical use and this may limit broad adoption.

Aim: This study was a first step to develop a shorter and/or more convenient version of the HJHS for the measurement of joint function in children and young adults with haemophilia, by combining real-life data and expert opinion.

Methods: A cross-sectional multicenter secondary analysis on pooled data of published studies using the HJHS (0-124, optimum score 0) in persons with haemophilia A/B aged 4-30 was performed. Least informative items, scoring options and/or joints were identified. An expert group of 19 international multidisciplinary experts evaluated the results and voted on suggestions for adaptations in a structured meeting (consensus set at ≥ 80%).

Results: Original data on 499 persons with haemophilia from 7 studies were evaluated. Median age was 15.0 years [range 4.0-29.9], 83.2% had severe haemophilia and 61.5% received prophylaxis. Median (IQR) HJHS total was 6.0 (1.0-17.0). The items 'duration swelling' and 'crepitus' were identified as clinically less informative and appointed as candidates for reduction.

Conclusion: Analysis of 499 children and young adults with haemophilia showed that the HJHS is able to discriminate between children and adults and different treatment regimens. Reduction of the items 'duration swelling' and 'crepitus' resulted in

[12]Department of Haematology, Oslo University Hospital, Oslo, Norway
[13]Hemophilia and Thrombosis Center, University of Colorado Anschutz Medical Campus, Aurora, CO, USA
[14]Department of PMR, Christian Medical College, Vellore, India
[15]Hemophilia of Georgia Center for Bleeding & Clotting Disorders of Emory, Emory University School of Medicine, Atlanta, GA, USA
[16]Department of Orthopaedics, Christian Medical College, Vellore, India
[17]CHU Sainte-Justine, Montreal, Canada
[18]Department of Radiology, Clementino Fraga Filho University Hospital, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
[19]Clinic of Children's diseases, Faculty of Medicine, Vilnius University, Vilnius, Lithuania
[20]Department of Rehabilitation, Paediatric Physical Therapy, Radboud university medical center, Nijmegen, The Netherlands

Correspondence
Isolde A.R. Kuijlaars, Van Creveldkliniek, University Medical Center Utrecht, Utrecht University, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands.
Email: i.a.r.kuijlaars-2@umcutrecht.nl

the $HJHS_{short}$, which had the same discriminative ability. Additional steps are needed to achieve a substantially shorter HJHS assessment.

KEYWORDS
expert opinion, haemophilia, joints, outcome measures, physical examination

1 | INTRODUCTION

Musculoskeletal assessment is an important component of the comprehensive care program for persons with haemophilia (PWH).[1] The Hemophilia Joint Health Score (HJHS) is recommended to evaluate joint health in clinical care and research.[2] The HJHS 1.0 was developed in 2003 and further developed in versions 2.0 and 2.1 by the International Prophylaxis Study Group (IPSG) Musculoskeletal Health Expert Working Group (EWG) for detection of early joint changes in children and youth with haemophilia.[3,4] Also in adolescents, the HJHS is a feasible and reliable tool.[5]

HJHS assessment including scoring takes approximately 45-60 minutes per patient, which has been felt, by some, to be impractical and infeasible, especially in a busy clinical setting. Therefore, there is a demand for a shorter version of the HJHS for clinical practice and time limited settings. With more studies using the HJHS being conducted internationally over the last 15 years, there is an opportunity to determine which items, scoring options and joints are universally important for different patient populations.

This study was a first step to develop a shorter and/or more convenient version of the HJHS for the measurement of joint function in children and young adults (aged 4-30 years) with haemophilia, by combining real-life data and expert opinion.

Key points

- The Hemophilia Joint Health Score (HJHS) is considered to be too time consuming.
- Real-life HJHS data and expert opinion were combined to develop a shorter HJHS.
- The items 'duration swelling' and 'crepitus' were identified as candidates for item reduction.
- Additional steps are needed to achieve a substantially more time efficient HJHS assessment.

2 | MATERIALS AND METHODS

This study was a multicenter secondary analysis of pooled data of published studies using the HJHS. After statistical analyses of the pooled data, results of this study were discussed and suggestions for adaptations were formally voted on, in an international expert meeting on 3 October 2019 in Utrecht (The Netherlands). This blended methodology was chosen since no criterion standard for the construct 'joint health' is available. In addition, consensus between HJHS developers, users and investigators is needed for adaptations of the HJHS, as well as implementation of recommended adaptations.

The Medical Research Ethical Committee (MREC) of the University Medical Center Utrecht reviewed and approved the study (17-499/C).

2.1 | Pooling of published HJHS data

A literature search identified 48 studies published between 2006 and 2019, which used either the HJHS 1.0, HJHS 2.0 or HJHS 2.1 in PWH. Forty-three of these studies had unique data. Inclusion criteria were PWH A (FVIII) or B (FIX) of all severities, aged 4-30 years. PWH were excluded if there were fewer than 5 complete items or fewer than 4 complete joints on the HJHS assessment. First, studies with <20 eligible PWH (n = 9), without full text papers (n = 1) or without HJHS data for all joints (n = 5) were excluded. The first validation study of the HJHS was also excluded.[4] Second, of the remaining 27 studies, the authors (IK, JN, BF, KF) selected studies to create a heterogeneous mix of PWH from different countries and treatment regimens, taking into account existing collaborations with authors of the studies like the IPSG Musculoskeletal Health EWG members. A sample size of 500 has been recommended for this sort of work.[6]

Authors of 16 papers were invited to share the original HJHS data (with scores on the item level) and patient characteristics. Two authors declined the invitation because HJHS item scores were unavailable. Seven authors did not reply to our request to share the data. Eventually, data of all children (4-17 years) and young adults (18-30 years) from seven remaining studies were included in the analysis.[7-13] One study used HJHS version 1.0,[8] one study used HJHS version 2.0[7] and the other studies used HJHS version 2.1.[9-13]

2.2 | Measurements

Patient characteristics collected for the included datasets were age at HJHS assessment, type of haemophilia (A or B), severity of the disease (mild [factor 0.06 IU/mL-0.40 IU/mL], moderate [factor 0.01 IU/mL- 0.05 IU/mL] or severe [factor <0.01 IU/mL]), clotting factor regimens (prophylaxis yes/no and start prophylaxis before age of 3 years yes/no) and current inhibitor status for each individual patient.

2.2.1 | HJHS

The HJHS 2.1 is the most recent version and consists of assessments of swelling (0-3), duration of swelling (0-1), muscle atrophy (0-2), crepitus on motion (0-2), flexion loss (0-3), extension loss (0-3), joint pain (0-2) and strength (0-4) for elbows, knees and ankles and a global gait score (0-4). Scores range from 0 to 20 per joint and the global gait score ranges from 0 to 4, resulting in a total HJHS score from 0 to 124 points.[14] A higher score indicates worse joint health. Scores of version HJHS 1.0 were converted to HJHS 2.1 by recoding of the original data for the items flexion loss, extension loss and gait (per joint) and deleting of the items axial alignment and instability. For datasets of version HJHS 2.0 and HJHS 2.1, scores on 'flexion loss' and 'extension loss' were copied.

2.3 | Statistical analyses

Patient characteristics were presented as proportions or medians (interquartile ranges [IQR:P25-P75]). Descriptive analyses (median [IQR], proportions of score categories) were performed for the HJHS total scores and joint scores. HJHS scores were compared for two age groups (children [4-17years] vs. adults [18-30 years]) and two different treatment regimens, defined as less intensive treatment (Romania, Pakistan, Lithuania, Brazil, USA) vs. intensive treatment (the Netherlands, UK, Canada) according to access to (early) prophylaxis (see Table 1).

To identify redundant items the following aspects were evaluated.

1. *Inter-item* correlations were evaluated. Inter-item correlations calculated with Spearman's rho <0.2 indicated items, which

TABLE 1 Patient characteristics

		Intensive treatment (n=220)	Less intensive treatment (n=279)
Children (n = 275)		**n=183**	**n=142**
	Age (years), median (IQR)	11.6 (8.9-14.7)	11.8 (9.0-15.0)
	Haemophilia severity, %		
	Mild	0	2.8
	Moderate	8.2	19.0
	Severe	91.8	78.2
	Prophylaxis, %	92.3	33.1
	Early prophylaxis (<3 y) / prophylaxis, %	67.8[a]	26.1[b]
Adults (n = 174)		**n=37**	**n=137**
	Age (years), median (IQR)	24.6 (20.9-27.2)	23.8 (20.9-26.8)
	Haemophilia severity, %		
	Mild	0	7.3
	Moderate	0	20.4
	Severe	100	72.3
	Prophylaxis, %	81.1	44.5
	Early prophylaxis (<3 y) / prophylaxis, %	14.3	0

[a]Missing n=20.

[b]Missing n=1.

do not correlate with any of the others and >0.9 indicated item redundancy.[15]
2. Component loadings on exploratory factor analyses were evaluated. Factor loadings <0.5 were considered indicators of item redundancy.[15] Model fit was evaluated with the Root Mean Square Error of Approximation [RMSEA].
3. Internal consistency calculated with Cronbach's α and internal consistency after item deletion were evaluated on joint level. Cronbach's α should be between 0.7 and 0.9; a higher Cronbach's α after item deletion was considered a reason to eliminate an item.[15] Global gait was included for the knees and ankles.
4. *Item-total* correlations for total joint scores were evaluated. Item-total correlations calculated with Spearman's rho <0.3 were indicators for an item that did not contribute to measurement of the construct.[15]
5. Proportions of zero and maximum scores on HJHS items were analysed for each joint (elbow, knee, ankle) to detect floor- and ceiling effects (≥85% zero or maximum scores on items),[4,15] in two age groups (children vs. adults) and two different treatment regimens, defined as less intensive treatment (Romania, Pakistan, Lithuania, Brazil, USA) vs. intensive treatment (the Netherlands, UK, Canada).

After item deletion a shortened HJHS total score ($HJHS_{short}$) was calculated. To evaluate the ability to discriminate between various patient groups, median HJHS total scores ($HJHS_{full}$ and $HJHS_{short}$) were calculated for children vs. adults and PWH receiving less intensive treatment vs. intensive treatment. In addition, proportions of PWH with affected joints were calculated with a cut-off point of ≥4 score for $HJHS_{full}$ and ≥3 score for $HJHS_{short}$, as HJHS scores up to 3 were shown in healthy subjects based on the items crepitus and flexion loss.[16]

For the comparison of the $HJHS_{full}$ and $HJHS_{short}$, scores were normalized from 0 to 100. Spearman's correlations, two-way mixed consistency Intraclass Correlation Coefficient (ICC) and Limits of Agreement (LoA) were calculated with the normalized scores.

To identify which scoring options were scored, endorsement for all scoring options for each item (% of options) was evaluated. To identify which joints were affected, descriptive analyses on joint level were performed.

2.4 | Expert meeting

Nineteen international experts participated in the 1-day expert meeting. A purposive sample of members of the IPSG Musculoskeletal Health EWG, experienced users of the HJHS and investigators of the included studies was selected. The expert group included eight physicians and eleven physical therapists. The expert meeting started with a presentation of the analysis of the pooled data and published literature on the HJHS scoring system,[17] crepitus in healthy subjects[16,18] and HJHS use for monitoring joint changes.[19] This was followed by three structured discussion sessions about the topics: item reduction, scoring options and number of joints assessed in the HJHS. Each discussion included five steps: presentation of results of the pooled data and statements for voting; questions for clarification of the results; first voting; discussion; and final voting. Experts voted anonymously with the online tool Mentimeter.com. Results of each vote were shown to the experts when all experts completed the vote on a statement. If at least 80% of the experts agreed, consensus was reached about a statement. BF moderated the discussion sessions.

3 | RESULTS

3.1 | Patient characteristics

In the seven studies, 499 PWH A or B (children [n = 325]; young adults [n = 174]) were included. Four PWH were excluded because they had <5 completed items or <4 completed joints assessed on the HJHS. The data are from Romania, the Netherlands, United Kingdom, Pakistan, Lithuania, Brazil, Canada and the United States. Patient characteristics according to age and treatment intensity are shown in Table 1. Median age at the time of HJHS assessment was 15.0 years (IQR 10.4-21.3, range 4.0-29.9). Most PWH had severe haemophilia (n = 415, 83.2%). More than half of the PWH (n = 307, 61.5%) used prophylaxis and 38.1% of these PWH had received early prophylaxis. Seventeen PWH (3%) had an inhibitor at HJHS assessment.

For the data from Pakistan about treatment regimen and from the United States about start of prophylaxis, assumptions were made, after contact with the authors and published patient characteristics.

3.2 | HJHS total and joint scores

The median (IQR) HJHS total score was 6.0 (1.0-17.0), with a range of 0-63. Twenty-one per cent of the PWH had a total score of 0 (children 26%;young adults 10%). Young adults had higher HJHS scores (11.5 [4.0-23.0]) than children (5.0 [0.0-12.0]). PWH receiving less intensive treatment showed higher HJHS scores (12.0 [5.0-26.0]) than PWH receiving intensive treatment (2.0 [0.0-7.0]). The ankles were the most affected joints, followed by the knees and elbows (see Figure 1).

3.3 | Discussion session 1: HJHS items

Figure 2 shows the process of discussion session 1 from the statistical analyses of the pooled data up to the validation of the $HJHS_{short}$.

3.3.1 | Selection of items eligible for item reduction

Reduction in item number was the first technique explored to reduce the time needed for HJHS assessment.

Inter-item correlations suggested no items were eligible for item reduction, since items did not show correlations >0.9 or <0.2.

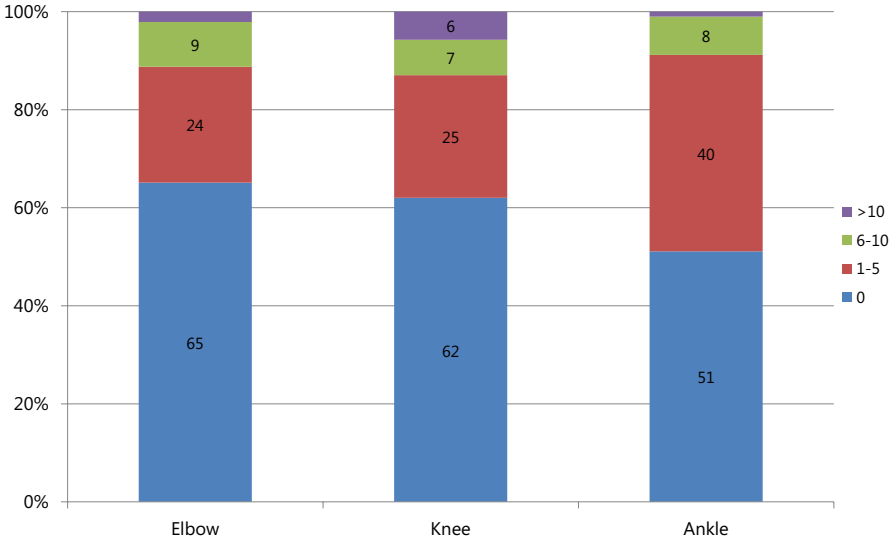

FIGURE 1 HJHS joint scores for all PWH (n = 499). HJHS joint scores > 10:2% for the elbow, 1% for the ankle [Colour figure can be viewed at wileyonlinelibrary.com]

Swelling' and 'duration swelling' showed the strongest correlation (r = 0.78-0.80) for elbows, knees and ankles.

The exploratory factor analyses suggested no items were eligible for item reduction. A 3-factor model was selected which included all HJHS items of the elbows, knees and ankles. Three factors were identified, namely elbows, knees and ankles. The model fit of the 3-factor model was good (RMSEA = 0.05). The highest factor loading for each item was >0.5. In addition, each joint was analysed separately using 1-factor models. The model fits of the 1-factor models were moderate (RMSEA = 0.07-0.08). For all items, the factor loadings were >0.5.

The internal consistency analyses suggested no items were eligible for item reduction. HJHS items were strongly related (Cronbach's α = 0.78-0.87) without a distinct increase in Cronbach's α after item deletion of separate items, except from item deletion of global gait in the knee joint.

In addition, *item-total* correlations showed high correlations (r = 0.37-0.69), thus identifying no candidates for item reduction.

Proportions of zero scores on HJHS items were analysed in four groups stratified by age and treatment intensity. The proportions of zero scores were highest for 'duration swelling' (varying from 78% to 98% for the different joints) and lowest for 'global gait' (35%-64%). The other items had proportions of zero scores of 63%-99%. Proportions of zero scores were higher in children and more intensively treated PWH.

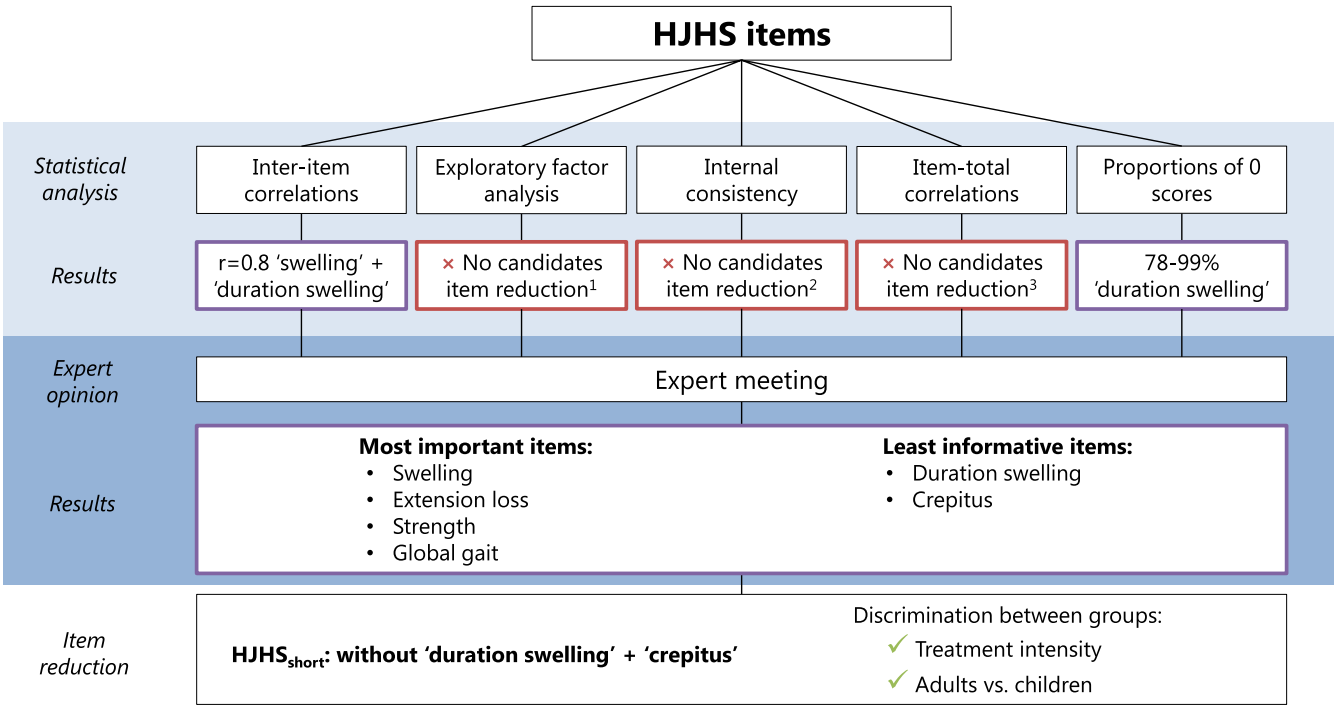

FIGURE 2 Flow chart of discussion session 1. Criteria for item reduction were: 1 Factor loadings < 0.5; 2 Cronbach's alpha < 0.7 or > 0.9; 3 correlations < 0.3 [Colour figure can be viewed at wileyonlinelibrary.com]

Proportions of zero scores are shown in Table 2. Detailed data of the 3-factor exploratory factor analysis and 1-factor exploratory analyses are shown in the Supporting information.

3.3.2 | Expert voting 1

The results of the voting are shown in Table 3. The experts reached consensus that 'duration swelling' (95%) and 'crepitus' (95%) are redundant items. Experts discussed the reliability of the item 'duration swelling' and the potential impact of recall bias on reliability, considering this is part of the clinical history rather than physical examination of the joint. For 'crepitus', an important argument for dropping it was that crepitus is also reported frequently in healthy people.[16,18] The items considered as most important were swelling (100%: important), extension loss (100%), strength (95%) and global gait (84%). In addition, the experts discussed the item 'global gait': whether it should be part of the HJHS as a tool assessing structure and function, or whether 'global gait' should be scored separately.

3.3.3 | Validation of $HJHS_{short}$

An $HJHS_{short}$ was created by deletion of the items 'duration swelling' and 'crepitus' (0-106). Abnormal HJHS joint scores based on 'crepitus' only were observed infrequently (elbows 1.0%; knees 3.6%; ankles 5.5%). Abnormal HJHS joint scores based on 'duration swelling' only were not observed. HJHS total scores for the $HJHS_{full}$ and the $HJHS_{short}$ and proportions of affected joints are shown in Table 4. The proportions of affected joints ($HJHS_{full} \geq 4$; $HJHS_{short} \geq 3$) were slightly higher for the $HJHS_{short}$. $HJHS_{short}$ was still able to discriminate between children vs. adults and PWH with less intensive treatment vs. intensive treatment. The normalized $HJHS_{full}$ and $HJHS_{short}$ correlated strongly in children and adults ($r = 0.98$) and in PWH with less ($r = 0.99$) and more intensive treatment ($r = 0.97$), with an ICC of 0.99 and LoA of −3.1 to 3.3 for the normalized scores.

3.4 | Discussion session 2: scoring options

3.4.1 | Frequency of endorsement for scoring options

Reduction of the number of scoring options for each item may be another way to reduce the time needed for HJHS assessment. Frequencies of endorsement for all scoring options are shown in Table 5. All items except 'duration swelling' and ' global gait' had scoring options, which were scored in ≤ 5% of the PWH.

TABLE 2 Proportions of zero scores in PWH with intensive treatment vs. less intensive treatment and children vs. adults.

		Intensive treatment (n=220)			Less intensive treatment (n=279)		
		Elbow	Knee	Ankle	Elbow	Knee	Ankle
Children (n=275)	% zero scores		n=183			n=142	
	Swelling	97	98	86	79	74	83
	Duration swelling	98	98	92	87	85	91
	Atrophy	97	95	90	79	66	75
	Crepitus	98	92	85	88	78	87
	Flexion loss	95	96	93	72	74	80
	Extension loss	94	97	88	84	84	88
	Pain	98	97	97	78	74	90
	Strength	98	98	92	70	67	75
	Global gait		64			35	
Adults (n=174)	% zero scores		n=37			n=137	
	Swelling	99	95	92	86	77	72
	Duration swelling	99	95	92	91	86	78
	Atrophy	89	89	78	86	71	73
	Crepitus	91	92	78	80	69	70
	Flexion loss	66	89	74	74	73	72
	Extension loss	68	93	70	75	86	76
	Pain	97	99	97	80	77	78
	Strength	93	95	81	82	71	74
	Global gait		57			40	

In grey: <85% zero scores.

TABLE 3 Results of discussion session 1 aimed at identifying redundant HJHS items

Item	Voting 1 Redundant, n (%)	Voting 2 Redundant, n (%)
Swelling	0 (0)	0 (0)
Duration swelling	11 (58)	**18 (95)**
Atrophy	7 (37)	8 (42)
Crepitus	13 (68)	**18 (95)**
Flexion loss	2 (11)	7 (37)
Extension loss	0 (0)	0 (0)
Pain	8 (42)	12 (63)
Strength	6 (33)[a]	1 (5)
Global gait	4 (21)	3 (16)

Note: Question to experts: Is this item important or redundant? Answer options: important/redundant.

[a]18 voters during first voting.

TABLE 4 Comparison of the $HJHS_{full}$ total score vs. the $HJHS_{short}$ total score, after item deletion of 'duration swelling' and 'crepitus'

	$HJHS_{full}$ (0-124)		$HJHS_{short}$ (0-106)	
	median (IQR)	% affected (≥4)	median (IQR)	% affected (≥3)
Intensive treatment (n = 220)	2.0 (0.0-7.0)	43.2	2.0 (0.0-5.0)	44.5
Less intensive treatment (n = 279)	12.0 (5.0-26.0)	78.5	10.0 (4.0-22.0)	80.3
Children (n = 325)	5.0 (0.0-12.0)	55.7	4.0 (0.0-10.0)	57.8
Adults (n = 174)	11.5 (4.0-23.0)	76.4	9.0 (3.0-20.0)	77.0

Note: Proportions of PWH with affected joint were calculated with a cut-off point ≥ 4 for $HJHS_{full}$ and ≥ 3 for $HJHS_{short}$, according HJHS scores from 0-3 shown in healthy subjects with scores on crepitus and flexion loss.[16]

3.4.2 | Expert voting 2

The results of the voting on scoring options are shown in the Supporting information. For 'pain', 79% of the experts voted that the scoring options could be reduced from three categories (*no pain through active range of motion/no pain through active range; only pain on gentle overpressure or palpation/pain through active range*) to a binominal scoring. An important argument against reduction of scoring options was that reducing the scoring options would only result in a minor reduction of the duration of HJHS assessment. It was decided that reduction of scoring options was not a feasible suggestion for shortening HJHS assessment.

3.5 | Discussion session 3: joints

3.5.1 | Joints

Reduction of the number of joints which needs assessment may be another way to reduce the time needed for HJHS assessment. The ankles were the most frequently affected joints, followed by the knees and elbows (see Figure 1). Ankles were most frequently affected in PWH on intensive treatment, while knees were most frequently affected in PWH on less intensive treatment.

3.5.2 | Expert voting 3

The results of the voting on joints are shown in the Supporting information, which proposed measuring a reduced number of joints. During the discussion, experts suggested that screening of joints instead of a full HJHS assessment could be a way to reduce time of HJHS assessment: 'assess all joints that fail a screening examination of medical/bleeding history and a physical examination'. However, a decision regarding which items to screen and how was considered beyond the scope of this meeting. Another topic discussed by the experts was that the most affected joint is not always the joint which needs the most attention. The experts reached consensus (94%) that for clinical practice a way to reduce assessment time is that only joints that fail a screening examination should be assessed with the full HJHS. The experts did not reach consensus (74%) about

TABLE 5 Distribution (%) of the scoring options of all HJHS items, for all PWH (n = *499*)

Scoring option	Swelling	Duration swelling	Atrophy	Crepitus	Flexion loss	Extension loss	Pain	Strength	Global gait
0	85.4	90.5	82.8	84.1	81.9	85.8	87.5	82.8	48.9
1	9.4	9.5	14.3	12.5	9.5	7.7	9.8	11.4	15.4
2	***4.7***		***2.9***	***3.4***	***4.5***	***3.4***	***2.7***	***2.9***	13.7
3	***0.4***				***4.0***	***3.0***		***1.8***	8.7
4								***1.1***	13.3

Note: In bold/italics: proportions < 5%.

the statements for a research setting, which proposed measuring a reduced number of joints.

4 | DISCUSSION

This study describes real-life HJHS data of 499 children and young adults with different treatment regimens combined with international expert opinion as a first step to develop a more convenient version of HJHS. The items 'duration swelling' and 'crepitus' were identified as candidates for item reduction. The resulting $HJHS_{short}$ was still able to discriminate between different ages and treatment regimens. Another way of shortening HJHS assessment for clinical practice suggested by the experts was a screening examination to select joints which need full HJHS assessment.

4.1 | Internal and external validity

This is the first study presenting HJHS data of 499 PWH from heterogeneous populations. The results showed different patterns of HJHS scoring in children vs. young adults and PWH with less intensive treatment vs. intensive treatment. We observed a wide variety of scores (range HJHS total: 0-63) with only 21% of the persons achieving a HJHS total score of 0 in this relatively young PWH with a wide range in treatment intensity. The real-life data in the present study were representative for clinical use and research purposes, while variation between raters was unavoidable in this study design. Some items show more variability in scoring between raters, which is a limitation of the HJHS, despite the good overall interobserver reliability.[20,21] In absence of a cut-off score for affected joints according the HJHS, a cut-off score was chosen above scores (0-3) shown in healthy adults which was established in a single observer study.[16]

In addition to the use of real-life data, international expert opinion of HJHS developers, clinical HJHS users and investigators using the HJHS increased the clinical value of the results. The experts who participated in the discussion sessions are representative of HJHS developers, users and researchers. This blended approach was used to compensate for the absence of a gold standard.

4.2 | Comparison with other studies

According to the analyses of the pooled data, 'crepitus' was not a candidate for item reduction. However, the experts voted that 'crepitus' could be regarded redundant because crepitus is a sign, which is also reported frequently (13%-14%) for the knees in healthy children and young adults.[16,18] Despite 'flexion loss' being reported in ankles of healthy young adults (12%)[16] and ROM assessment is time consuming, only 37% of the experts voted that this item was redundant and should be eliminated.

Furthermore, the experts voted that 'duration swelling' was redundant. According to the analyses of the pooled data, 'duration swelling' was indeed a candidate for item reduction based on the floor effects. An additional argument of the experts was potential recall bias, which could lower the reliability of this item. Although interobserver and test-retest reliability of this item were reported in two studies (ICC = 0.44-0.90),[20,21] these findings do not support the experts' argument of recall bias since these studies did not address the risk on recall bias over six months.

4.3 | Clinical implications and future research

The items 'duration swelling' and 'crepitus' were identified as candidates for item reduction. Dropping these two items will not lead to a substantial gain in time. Therefore, it is relevant to search for further ways to achieve shorter joint assessment in clinical practice. As suggested by the experts, a next step to explore is joint screening to select the joints which need full assessment.

Besides shortening the HJHS to make joint assessment more feasible in routine clinical practice, additional focus on standardization of items is needed.

5 | CONCLUSION

This study in 499 PWH showed that the HJHS is able to discriminate between children and adults and different treatment regimens. Based on expert (n = 19) consensus, reduction of the items 'duration swelling' and 'crepitus' resulted in the $HJHS_{short}$, which had the same discriminative ability. To achieve a shorter joint assessment in clinical practice, joint screening to select the joints which need full assessment was suggested.

ACKNOWLEDGEMENTS

The authors would like to thank A Willemsen and O Versloot for their help in the organization of the expert meeting.

CONFLICT OF INTEREST

J van der Net, BM Feldman, SM Funk, P Hilliard, M Manco-Johnson and P Petrini are co-inventors of the Haemophilia Joint Health Score (HJHS). CL Kempton has received honoraria for consulting and participating in advisory boards from Spark Therapeutics, Octapharma, Genetech, and Pfizer and research support from Novo Nordisk. MJ Manco-Johnson has received honoraria for Advisory Boards from BioMarin, BioVerativ, CSL Behring, Novo Nordisk, Sparks and Takeda. P Petrini has acted as a paid consultant to Roche, Baxalta, Sobi and Baxter. The other authors have no competing interests.

AUTHOR CONTRIBUTION

IAR Kuijlaars, J van der Net, BM Feldman and K Fischer contributed to the design of the study; IAR Kuijlaars performed the statistical analyses; IAR Kuijlaars wrote the first draft of the paper; all authors contributed to interpretation of the data, modification of statistical analyses and the writing of the manuscript.

ORCID

Isolde A.R. Kuijlaars https://orcid.org/0000-0003-2920-2258
Janjaap van der Net https://orcid.org/0000-0003-2606-5104
Brian M. Feldman https://orcid.org/0000-0002-7813-9665
Magnus Aspdahl https://orcid.org/0000-0002-4427-5159
Melanie Bladen https://orcid.org/0000-0002-7413-1906
Rubén Cuesta-Barriuso https://orcid.org/0000-0002-4276-6982
Ruth E.D. Matlary https://orcid.org/0000-0002-3961-0207
Christine L. Kempton https://orcid.org/0000-0002-2391-2016
Marilyn Manco-Johnson https://orcid.org/0000-0001-5215-1482
Pradeep Poonnoose https://orcid.org/0000-0001-7715-9982
Sylvia Thomas https://orcid.org/0000-0002-2375-8602
Merel A. Timmer https://orcid.org/0000-0003-1910-999X
Sonata Saulyte Trakymiene https://orcid.org/0000-0003-3576-7116
Kathelijn Fischer https://orcid.org/0000-0001-7126-6613

REFERENCES

1. Srivastava A, Brewer AK, Mauser-Bunschoten EP, et al. Guidelines for the management of hemophilia. *Haemophilia*. 2013;19.
2. Fischer K, Poonnoose P, Dunn AL, et al. Choosing outcome assessment tools in haemophilia care and research: A multidisciplinary perspective. *Haemophilia*. 2016;23(1):1-14.
3. Feldman BM, Funk S, Lundin B, et al. Musculoskeletal measurement tools from the International Prophylaxis Study Group (IPSG). *Haemophilia*. 2008;14:162-169.
4. Feldman BM, Funk SM, Bergstrom B-MM, et al. Validation of a new pediatric joint scoring system from the international hemophilia prophylaxis study group: Validity of the hemophilia joint health score. *Arthritis Care Res*. 2011;63:223-230.
5. Fischer K, de Kleijn P. Using the Haemophilia Joint Health Score for assessment of teenagers and young adults: Exploring reliability and validity. *Haemophilia*. 2013;19:944-950.
6. Mokkink LB, Terwee CB, Patrick DL, et al.COSMIN checklist manual 2012.
7. Groen W, Van der Net J, Lacatusu AM, et al. Functional limitations in Romanian children with haemophilia: Further testing of psychometric properties of the Paediatric Haemophilia Activities List. *Haemophilia*. 2013;19:116-125.
8. Fischer K, Carlsson KS, Petrini P, et al. Intermediate-dose versus high-dose prophylaxis for severe hemophilia: Comparing outcome and costs since the 1970s. *Blood*. 1970s;122:1129-1136.
9. Nijdam A, Bladen M, Hubert N, et al. Using routine Haemophilia Joint Health Score for international comparisons of haemophilia outcome: Standardization is needed. *Haemophilia*. 2016;22:142-147.
10. Khanum F, Bowen DJ, Kerr BC, Collins PW. Joint health scores in a haemophilia A cohort from Pakistan with minimal or no access to factor VIII concentrate: correlation with thrombin generation and underlying mutation. *Haemophilia*. 2014;20:426-434.
11. Carneiro JDA, Blanchette V, Ozelo MC, et al. Comparing the burden of illness of haemophilia between resource-constrained and unconstrained countries: the São Paulo-Toronto Hemophilia Study. *Haemophilia*. 2017;23:682-688.
12. Kempton CL, Recht M, Neff A, et al. Impact of pain and functional impairment in US adults with haemophilia: Patient-reported outcomes and musculoskeletal evaluation in the pain, functional impairment and quality of life (P-FiQ) study. *Haemophilia*. 2018;24:261-270.
13. Saulyte-Trakymiene S, Juodyte A, Jusinskaite V, Kulikauskaite R. Systematic evaluation of hemophilic arthropathy in Lithuania. *J Med Sci*. 2019;7:1-15.
14. International Prophylaxis Study Group. Hemophilia Joint Health Score 2.1 Instruction Manual n.d.
15. De Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A practical guide*. New York: Cambridge University Press; 2011.
16. Sluiter D, Foppen W, de Kleijn P, Fischer K. Haemophilia Joint Health Score in healthy adults playing sports. *Haemophilia*. 2014;20:282-286.
17. Ribeiro T, Abad A, Feldman BM. Developing a new scoring scheme for the Hemophilia Joint Health Score 2.1. Res Pract. *Thromb Haemost*. 2019;3:405-411.
18. Hacker MR, Funk SM, Manco-Johnson MJ. The Colorado Haemophilia Paediatric Joint Physical Examination Scale: Normal values and interrater reliability. *Haemophilia*. 2007;13:71-78.
19. Kuijlaars IAR, Timmer MA, de Kleijn P, Pisters MF, Fischer K. Monitoring joint health in haemophilia: Factors associated with deterioration. *Haemophilia*. 2017;23(6):934-940.
20. Sun J, Hilliard PE, Feldman BM, et al. Chinese Hemophilia Joint Health Score 2.1 reliability study. *Haemophilia*. 2014;20:435-440.
21. Hilliard P, Funk S, Zourikins N, et al. Hemophilia joint health score reliability study. *Haemophilia*. 2006;12:518-525.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Kuijlaars IAR, van der Net J, Feldman BM, et al. Evaluating international Haemophilia Joint Health Score (HJHS) results combined with expert opinion: Options for a shorter HJHS. *Haemophilia*. 2020;26:1072–1080. https://doi.org/10.1111/hae.14180