

BRIEF REPORT

Developing a new scoring scheme for the Hemophilia Joint Health Score 2.1

Tiago Ribeiro BMSc¹ | Audrey Abad BSc² | Brian M. Feldman MD, MSc, FRCPC^{2,3,4}

¹Schulich School of Medicine & Dentistry, Western University, London, Ontario, Canada

²Child Health Evaluative Sciences Program, Research Institute, The Hospital for Sick Children, Toronto, Ontario, Canada

³Department 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

⁴Division of Rheumatology, The Hospital for Sick Children, Toronto, Ontario, Canada

Correspondence

Brian Feldman, Department Pediatrics, Division of Rheumatology, The Hospital for Sick Children, 555 University Ave., Toronto, ON M5G 1X8 Canada.

Email: brian.feldman@sickkids.ca

Abstract

Background: The Hemophilia Joint Health Score (HJHS) is a validated outcome tool developed for the assessment of joint health in people with hemophilia. The ordinal joint score assesses 9 items in 6 index joints. It is recognized as an optimal measurement of arthropathy in children and young adults. The aim of this study was to develop an updated scoring system for the HJHS that may overcome the limitations of its current ordinal scoring structure.

Methods: A survey was developed using 1000Minds decision-making software. Respondents were provided with discrete choice tasks of ranking alternatives to determine the preference weight, or relative importance, placed on different criteria for each HJHS item. The survey was distributed to an anonymous sample of health care professionals with extensive experience in the physical examination of joints in people with hemophilia.

Results: A total of 64 musculoskeletal health care professionals participated; with a 64% survey completion rate. The HJHS item weights provide a sum to 1.0; the highest-ranked item was extension loss (0.139) followed by swelling (0.121), whereas the lowest was duration of swelling (0.057) followed by muscle atrophy (0.08). Compared to the original, the relative efficiency of the new score was 5.4.

Conclusions: Observed differences in preference weights for HJHS items highlight the potential under- or overestimation of true joint health using the current ordinal scoring system. An updated scoring system using weighted items may improve the precision of HJHS assessment, leading to improved clinical management of joint health, while providing a robust research tool.

Essentials

- The Hemophilia Joint Health Score (HJHS) is an ordinal validated joint outcome assessment tool.
- A system using a discrete choice experiment may provide improvements to remove the current limitations of the HJHS score.
- An integrated approach may improve the HJHS for assessing joint health in clinical and research settings.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2019 The Authors. *Research and Practice in Thrombosis and Haemostasis* published by Wiley Periodicals, Inc on behalf of International Society on Thrombosis and Haemostasis.

1 | BACKGROUND

Hemophilia is a hereditary bleeding disorder, in which recurrent hemarthrosis may lead to rapidly progressive arthropathy.^{1,2} This may result in lasting functional impairments and participation restriction, as well as pain and deformity.^{3,4} People with hemophilia may have a reduced quality of life as a result of the physical and psychosocial impact of the disease.⁵

Musculoskeletal outcomes of patients with hemophilia are of great clinical importance.

According to the International Classification of Functioning, Disability and Health, “structure and function” is a major component of an individual's outcome in health and disability. Surveillance of musculoskeletal changes is recognized to be essential for timely patient evaluation and subsequent optimization of management.^{6–8}

The early initiation of prophylaxis has resulted in the need for a more sensitive assessment tool to identify the subtle joint changes that may lead to arthropathy.^{9,10} Therefore, a disease-specific tool focused on commonly affected joints would be more optimal than a generic musculoskeletal assessment tool.

There is currently no gold standard for the (latent) construct of “joint health.” The Hemophilia Joint Health Score (HJHS) was developed to assess joint health in people with hemophilia. Through several assessments of joint structure and function—collectively referred to as items—the HJHS produces a measure of joint health. The 9 items assessed to evaluate the status of a joint include swelling, duration of swelling, muscle atrophy, crepitus of motion, range of motion (extension and flexion loss), joint pain, strength, and gait. The summation of the 9 HJHS items in 6 index joints provides users with a relative indicator of joint health, with a lower HJHS representing superior joint health. HJHS scores may assist in the development of individualized musculoskeletal treatment plans or programs. The HJHS is an internationally validated physical examination assessment tool with excellent interobserver and test-retest reliability.^{9,10} In addition, the HJHS is frequently used in clinical studies and is thought to be optimal for assessing mild/moderate arthropathy in children and young adults.⁷

The HJHS, as currently scored, provides its users with ordinal data that are, perhaps incorrectly, treated as though they are numerical. For example, single-ordered category increase, or decrease in 2 HJHS item scores, will score identically with the current system but may not capture differences in the value of true joint health. Consequently, the clinical and research utility of the HJHS may be limited due to its ordinal structure.

Conjoint analysis can be used to develop a weighting scheme for measures such as the HJHS, addressing the limitations associated with simply summing individual ordinal attribute levels.¹¹ Discrete choice experiments are conjoint analyses; they are survey-based techniques that enable respondents to make repeated judgments on pairwise comparisons of attributes. The relative importance, or weight, placed on each attribute can then be estimated.¹¹ Applied

to the HJHS, this means the relative importance, or weight, respondents place on each item (and each level within each item) can be determined. This preference weighting can then be used to develop a scoring system that produces continuous scoring (weighted score) that gives more weight to the items considered to be of more relative importance for optimal joint health.

The purpose of this initiative is to use an adaptive, partial-profile, discrete choice experiment to lay the groundwork toward the development of an updated scoring system for the HJHS. Our goal is to transform the ordinal data created by the HJHS to continuous weighted scores determined by the importance respondents place on each item and level, addressing potential limitations of its current ordinal structure.

2 | METHODS

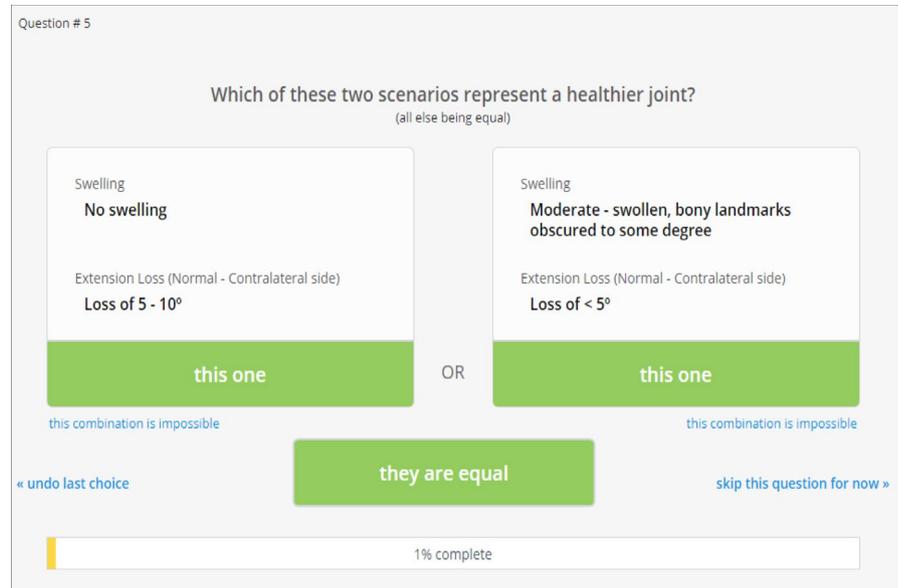
A conjoint analysis to determine the relative importance respondents place on each HJHS item was used. Our survey provided respondents with discrete choice tasks of ranking alternatives, which were then analyzed to provide the preference weight, or relative importance, individuals place on different criteria.

The survey was developed using 1000Minds decision-making software (www.1000minds.com). The specific method used by 1000Minds is known as the Potentially All Pairwise Rankings of all Possible Alternatives (PAPRIKA) method.¹² The PAPRIKA method involves a simple ranking measurement of decision makers' preferences rather than a scaling or ratio measurement. The process minimizes the number of pairwise comparisons respondents need to make by implicitly ranking items against each other using the data obtained from explicit pairwise comparisons. Preference values (weights) are generated for each individual participant using hypothetical real-world scenarios in contrast to other methods that solely produce aggregate data, which is a major advantage of the PAPRIKA method.¹² As such, our study was an adaptive, partial-profile, discrete-choice experiment.

The survey was distributed to an anonymous, random sample (random.org) of members of the Canadian Physiotherapists in Haemophilia Care, the World Federation of Haemophilia Musculoskeletal (WFH MSK) Committee, and the European Association for Haemophilia and Allied Disorders Physiotherapists Committee. These organizations were selected due to the members' extensive experience in the physical examination of joints and bleeding disorders. The respondents were predominantly pediatric and adult physiotherapists, with additional musculoskeletal experts from the WFH MSK committee. Due to the anonymity of the sample, no demographic information is available to report. All potential respondents were initially contacted via email, with subsequent reminders by email correspondence and a direct phone call.

The survey asked respondents to make repeated judgments between 2 hypothetical scenarios they believed represented a “healthier joint.” Each scenario consisted of a hypothetical joint scored by 2 HJHS items, as shown in Figure 1. Through repeated direct pairwise

FIGURE 1 Example of a decision scenario used in the 1000Minds Survey



comparisons of 2 scenarios, the preference weight for each item and each level was calculated by regression analysis.¹² The model contains 9 criteria with between 2 and 5 levels. This estimates an average total of 86 400 hypothetical individual simulations, resulting in approximately 93 decisions to be completed by each participant.

The mean, median, and standard deviation of the relative importance of each item was reported as a percentage; the sum of each item's relative importance (weight) is therefore 100%. As opposed to the original HJHS, higher scores in the weighted system indicate healthier joints.

The 1000Minds program performed random consistency checks whereby the survey participants were asked to explicitly rank a specific pairwise comparison they previously completed. Four consistency checks were performed for each respondent throughout the survey to assess responder reliability. Respondents with ≥ 2 valid consistency checks were included in an additional subgroup analysis.

To compare the relative efficiency of the original HJHS to the weighted HJHS, we applied the new weighted scoring system to the HJHS total scores from published data included in a study by Carneiro et al,¹³ comparing patients with hemophilia from both resource-constrained and -unconstrained countries. Independent samples *t*-tests were completed using both the original and weighted HJHS total scores.

3 | RESULTS AND DISCUSSION

A total of 64 individuals were contacted to participate in our survey with 41 (64%) participants completing the entire survey. The remaining 19 participants completed an average of 5.8% of the required decisions but declined to participate further. Consistency check results are found in Table 1. Subgroup analysis removed 3 respondents who showed inconsistency. This did not affect the overall results, leading to greater confidence in our total sample.

The order of the single levels within each item of the HJHS were predetermined in accordance with the HJHS scoring system; the item level representing a healthy joint was given a score of 100, and the worst possible level was given a score of 0. Regression analysis was used to provide the relative preference values for the 9 items of the HJHS as reported in Table 2. Normalized HJHS item weights sum to 1.0, and the mean scores for each single item level within each HJHS item are reported in Table 3. The HJHS item with the highest weight was extension loss (normal–contralateral side) (0.139), followed by swelling (0.121). The HJHS item with the lowest weight was duration of swelling (0.057), followed by muscle atrophy (0.08). The range in weights for different HJHS items demonstrate the large variability in the perceived impact these items have on determining joint health. This further strengthens the need for a more descriptive tool that accounts for the perceived impact these clinical indicators have on joint health.

Figure 2 illustrates the comparison of the original HJHS to the newly developed weighted HJHS. Results from the independent samples *t*-test indicated that the weighted HJHS, $t(98) = 7.748$; $P = 8.723e^{-12}$ had a significantly larger *T* value and a smaller *P* value than the original HJHS, $t(98) = 3.333$; $P = 0.001$ (relative efficiency = 5.4). The weighted HJHS has superior descriptive qualities,

TABLE 1 Summary of consistency checks performed in the 41 survey participants

Number of identical answers per repeated question (maximum = 4)	Number of participants (total = 41)
1	3
2	21
3	11
4	6

TABLE 2 Relative preference value for item level scores within each HJHS item reported as percent mean, median, and standard deviation (SD) of total responses generated by the regression model

HJHS item	Median (%)	Mean (%)	SD (%)
Swelling			
Severe—very swollen, no bony landmarks are visible	0	0	0
Moderate—swollen, bony landmarks obscured to some degree	2.7	4.5	3.4
Mild—swelling looks slightly “puffy”; all bony landmarks are visible	6.4	7.9	5.0
No swelling	11.2	12.1	6.1
Duration of swelling			
≥6 months	0	0	0
No swelling or <6 months	4.2	5.7	4.7
Muscle atrophy			
Severe—flattening of muscle belly is noted	0	0	0
Mild—mild flattening of muscle belly is noted	3.6	4.8	3.3
None	7.7	8.0	4.2
Crepitus of motion			
Severe—audible and/or palpable grinding and crunching during joint motion	0	0	0
Mild—slightly audible and/or palpable grinding and crunching during joint motion	4.6	6.6	5.4
None	10.0	10.9	6.7
Strength			
Trace or no muscle contraction	0	0	0
Able to partially complete range of motion against gravity	3.5	4.7	3.5
Holds test position with minimal resistance	6.8	7.8	4.9
Holds test position against gravity with moderate resistance	9.7	10.3	5.8
Holds test position against gravity with maximum resistance	11.6	13.0	7.1
Joint pain			
Pain through active range	0	0	0
No pain through active range; pain only on gentle overpressure or palpation	6.8	7.6	3.8
No pain through active range of motion	10.2	11.0	4.5
Global gait (walking, stairs, running, hopping on 1 leg)			
No skills are within normal limits	0	0	0
3 skills are not within normal limits	3.1	3.7	3.5
2 skills are not within normal limits	6.5	7.0	4.6
1 skill is not within normal limits	9.2	10.1	5.8
All skills are within normal limits	11.8	13.8	7.5
Flexion loss (normal–contralateral side)			
Loss of >20°	0	0	0
Loss of 11°–20°	3.3	4.0	2.9
Loss of 5°–10°	7.4	8.2	5.0
Loss of <5°	10.9	11.5	6.0
Extension loss (normal–contralateral side)			
Loss of >20°	0	0	0
Loss of 11°–20°	4.5	5.3	4.3
Loss of 5°–10°	8.7	10.0	6.4
Loss of <5°	11.4	13.9	7.6

Note: Values indicate how favorable a level was relative to other levels within an item; a higher score indicates it is more favorable when identifying a healthy joint (N = 41).

TABLE 3 Normalized HJHS item weights and item level scores (means)

HJHS item	Item weight (sum to 1)	Item levels	Item-level score (0-100)
Swelling	0.121	Severe—very swollen, no bony landmarks are visible	0
		Moderate—swollen, bony landmarks obscured to some degree	36.8
		Mild—swelling looks slightly “puffy”; all bony landmarks are visible	65
		No swelling	100
Duration of swelling	0.057	≥6 months	0
		No swelling or <6 months	100
Muscle atrophy	0.08	Severe—flattening of muscle belly is noted	0
		Mild—mild flattening of muscle belly is noted	60.5
		None	100
Crepitus of motion	0.109	Severe—audible and/or palpable grinding and crunching during joint motion	0
		Mild—slightly audible and/or palpable grinding and crunching during joint motion	59.9
		None	100
Strength	0.13	Trace or no muscle contraction	0
		Able to partially complete range of motion against gravity	36.2
		Holds test position with minimal resistance	60.5
		Holds test position against gravity with moderate resistance	79.3
		Holds test position against gravity with maximum resistance	100
Joint pain	0.11	Pain through active range	0
		No pain through active range; pain only on gentle overpressure or palpation	68.7
		No pain through active range of motion	100
Global gait (walking, stairs, running, hopping on 1 leg)	0.138	No skills are within normal limits	0
		3 skills are not within normal limits	26.6
		2 skills are not within normal limits	50.6
		1 skill is not within normal limits	73.3
		All skills are within normal limits	100
Flexion loss (normal-contralateral side)	0.115	Loss of >20°	0
		Loss of 11°-20°	35.2
		Loss of 5°-10°	71.4
		Loss of <5°	100
Extension loss (normal-contralateral side)	0.139	Loss of >20°	0
		Loss of 11°-20°	38.3
		Loss of 5°-10°	72
		Loss of <5°	100
Total	1.00		

Note: Item weights indicate how important each item is relative to other items when identifying a healthy joint (N = 41). Attributes for each item level are ranked highest (eg, healthiest) to lowest.

which may allow for smaller sample sizes in future studies. This updated system may not only result in a more efficient research tool but may also improve its use clinically.

The method of conjoint analysis using 1000Minds software was successfully implemented in rheumatology to determine preferences for outcome domains in gout¹⁴ and remission and response criteria development.¹⁵ Additionally, it has been used in preferences toward treatment,^{16,17} and inhibitor care in hemophilia patients.¹⁸ In all examples, the weights were determined by the preferences of the

health care professionals or patients/parents, supporting its applicability to this area.

Our sample was a random, representative group of hemophilia musculoskeletal experts, predominantly physiotherapists. We believe the groups we selected have the most valuable perspective given their daily use and application of musculoskeletal examination for people with hemophilia. However, perspectives of other members of the health care team and people with hemophilia will be important for future work. We recognize the importance of including

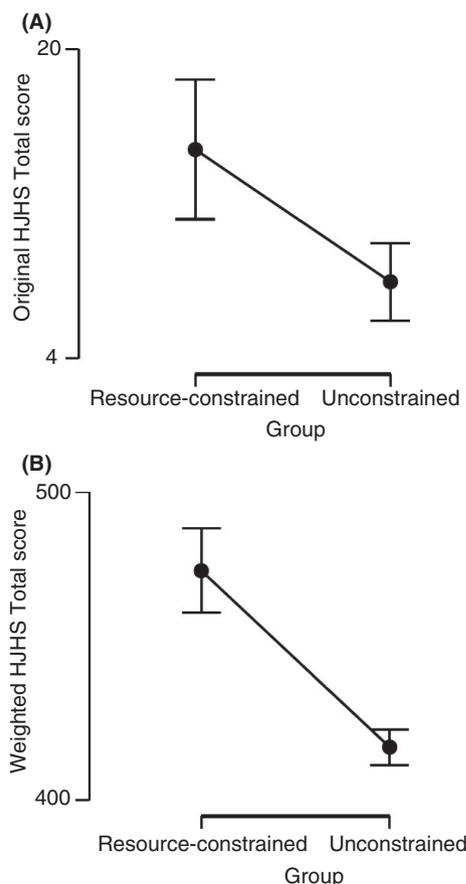


FIGURE 2 This figure illustrates the difference between (A) original HJHS total scores and (B) weighted HJHS total scores

patient preferences and plan to implement this information in a subsequent step toward its future development.

Due to the nature of conjoint analyses, some limitations were identified. Responder fatigue played a strong role in limiting the completion rate. Although the survey was adaptive, reducing the number of explicit comparisons made, the high number of possible pairwise comparisons led to an average of 73 decisions per completed survey. Our completion rate was adequate, but future work should consider methods to reduce respondent burden.

This activity is a preliminary step toward the development of an updated scoring system for the HJHS. As demonstrated by our analysis, a weighted HJHS scoring system should increase its value and utility. Moving forward, optimization of the HJHS should include the integration of patient and clinician preferences, leading to improved clinical assessment and management of joint health, while providing a more robust tool for research.

ACKNOWLEDGMENTS

We thank Sharon Funk, Pamela Hilliard, Marilyn Manco-Johnson, Pia Petrini, Jean St-Louis, Janjaap van der Net, and Nichan Zourikian (members of the Physical Health and Joint Function Expert Working Group of the International Prophylaxis Study Group) for their advice and support. A special thank you to Pamela Hilliard for her assistance with data collection.

RELATIONSHIP DISCLOSURE

Dr. Brian Feldman has a patent for the Hemophilia Joint Health Score 2.1 © with royalties paid to The Hospital for Sick Children, Centre Hospitalier Universitaire Sainte Justine, the Regents of the University of Colorado, Karolinska Hospital, University Medical Center Utrecht, 2009. Under license by The Hospital for Sick Children. All other authors do not have any disclosures.

AUTHOR CONTRIBUTIONS

TR assisted in the data collection and interpretation and analysis of the data, and wrote the first draft of the manuscript. AA coordinated the data collection, assisted in interpretation and analysis of the data, and critically revised the manuscript. BMF developed the concept and design of the project and assisted in the interpretation and analysis of the data and critically revised the manuscript.

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. Van den Berg HM, Dunn A, Fischer K, Blanchette VS. Prevention and treatment of musculoskeletal disease in the haemophilia population: role of prophylaxis and synovectomy. *Haemophilia*. 2006;12(Suppl 3):159-68.
3. Manco-Johnson MJ, Abshire TC, Shapiro AD, Riske B, Hacker MR, Kilcoyne R, et al. Prophylaxis versus episodic treatment to prevent joint disease in boys with severe hemophilia. *N Engl J Med*. 2007;357:535-44.
4. Valentino LA. Blood-induced joint disease: the pathophysiology of hemophilic arthropathy. *J Thromb Haemost*. 2010;8:1895-902.
5. Chen SL. Economic costs of hemophilia and the impact of prophylactic treatment on patient management. *Am J Manag Care*. 2016;22:s126-33.
6. Beeton K. Evaluation of outcome of care in patients with haemophilia. *Haemophilia*. 2002;8:428-34.
7. Blanchette VS, O'Mahony B, McJames L, Mahlangu JN. Assessment of outcomes. *Haemophilia*. 2014;20(Suppl 4):114-20.
8. Kuijlaars IAR, Timmer MA, de Kleijn P, Pisters MF, Fischer K. Monitoring joint health in haemophilia: factors associated with deterioration. *Haemophilia*. 2017;23:934-40.
9. Feldman BM, Funk SM, Bergstrom BM, Zourikian N, Hilliard P, van der Net J, 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 (Hoboken)*. 2011;63:223-30.
10. Hilliard P, Funk S, Zourikian N, Bergstrom BM, Bradley CS, McLimont M, et al. Hemophilia joint health score reliability study. *Haemophilia*. 2006;12:518-25.
11. Avila ML, Stinson J, Kiss A, Brandao LR, Uleryk E, Feldman BM. A critical review of scoring options for clinical measurement tools. *BMC Res Notes*. 2015;8:612.
12. Hansen P, Omblor F. A New method for scoring additive multi-attribute value models using pairwise rankings of alternatives. *J Multi-Crit Decis Anal*. 2008;15:87-107.
13. Carneiro JDA, Blanchette V, Ozelo MC, Antunes SV, Villaca PR, Young NL, et al. ; São Paulo-Toronto Hemophilia Study Group. Comparing the burden of illness of haemophilia between

- resource-constrained and unconstrained countries: the Sao Paulo-Toronto Hemophilia Study. *Haemophilia*. 2017;23:682–8.
14. Taylor WJ, Brown M, Aati O, Weatherall M, Dalbeth N. Do patient preferences for core outcome domains for Chronic Gout Studies support the validity of composite response criteria? *Arthritis Care Res (Hoboken)*. 2013;65:1259–64.
 15. de Lautour H, Taylor WJ, Adebajo A, Alten R, Burgos-Vargas R, Chapman P, et al. Development of preliminary remission criteria for gout using Delphi and 1000Minds consensus exercises. *Arthritis Care Res (Hoboken)*. 2016;68:667–72.
 16. Carlsson KS, Andersson E, Berntorp E. Preference-based valuation of treatment attributes in haemophilia A using web survey. *Haemophilia*. 2017;23:894–903.
 17. Furlan R, Krishnan S, Vietri J. Patient and parent preferences for characteristics of prophylactic treatment in hemophilia. *Patient Prefer Adherence*. 2015;9:1687–94.
 18. Lee WC, Joshi AV, Woolford S, Sumner M, Brown M, Hadker N, et al. Physicians' preferences towards coagulation factor concentrates in the treatment of haemophilia with inhibitors: a discrete choice experiment. *Haemophilia*. 2008;14:454–65.

How to cite this article: 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. <https://doi.org/10.1002/rth2.12212>