Received: 13 June 2022 | Revised: 13 November 2022 | Accepted: 16 November 2022

DOI: 10.1111/hae.14705

Haemophilia WFH WILEY

ORIGINAL ARTICLE

ACTIVLIM-Hemo: A new self-reported, unidimensional and linear measure of activity limitations in persons with haemophilia

Sébastien Lobet[1,2,3] | Catherine Lambert[1] | Anthe Foubert[4,5,6] | Valérie-Anne Chantrain[4,5,6] | Nathalie Roussel[4] | Mira Meeus[4,5,7] | Ann Devos[8] | Philip Maes[8] | Cedric Hermans[1] | Massimo Penta[9,10]

[1]Haemostasis and Thrombosis Unit, Division of Hematology, Clinique Universitaires Saint-Luc, Université Catholique de Louvain (UCLouvain), Brussels, Belgium

[2]Secteur de Kinésithérapie, Cliniques Universitaires Saint-Luc, Université Catholique de Louvain (UCLouvain), Brussels, Belgium

[3]Neuromusculoskeletal Lab (NMSK), Secteur des Sciences de la Santé, Institut de Recherche Expérimentale et Clinique, Université Catholique de Louvain (UCLouvain), Brussels, Belgium

[4]Research Group MOVANT, Department of Rehabilitation Sciences and Physiotherapy (REVAKI), University of Antwerp, Wilrijk, Belgium

[5]Pain in Motion, International Research Group

[6]Faculté des Sciences de la Motricité, Université Catholique de Louvain (UCLouvain), Louvain-la-Neuve, Belgium

[7]Department of Rehabilitation Sciences, Ghent University, Ghent, Belgium

[8]Hemostasis and Thrombosis Unit, Division of Pediatric Hematology, University Hospital Antwerp, Antwerp, Belgium

[9]Institute of Neuroscience, Université catholique de Louvain, Louvain-la-Neuve, Belgium

[10]Arsalis SRL, Glabais, Belgium

Correspondence
Sébastien Lobet, Haemostasis and Thrombosis Unit Division of Haematology, Cliniques Universitaires Saint-Luc, Avenue Hippocrate 10, BE-1200 Brussels, Belgium.
Email: sebastien.lobet@saintluc.uclouvain.be

Funding information
Pfizer Inc., Grant/Award Number: 55563735

Abstract

Introduction: To assess activity limitations in people with haemophilia (PwH), the self-reported Haemophilia Activity List (HAL) is widely employed, despite several methodological limitations impacting the interpretation of categorical scores. Modern psychometric approaches avoid these limitations by using a probabilistic model, such as the Rasch model. The ACTIVLIM is a Rasch-built measurement of activity limitations previously validated in several clinical conditions like neuromuscular disorders.

Aims: This study sought to develop the ACTIVLIM-Hemo, meaning an ACTIVLIM scale version specifically adapted to assess daily activity limitations in adult PwH.

Methods: Daily activities were assessed as "impossible," "difficult" or "easy" by 114 PwH (median age of 44 years) with 63 of them reassessed after 12 days. The Rasch Rating Scale model was used to identify activities delineating a unidimensional and linear scale unbiased by demographic and clinical status. Concurrent validity was determined through correlation with the HAL sub-scores and sum score.

Results: The ACTIVLIM-Hemo included 22 pertinent activities, with difficulties independent of demographic and clinical conditions, allowing a reliable measure of activity

limitations (PSI = .92) expressed on a linear and unidimensional scale in PwH (7%–100 % range, ceiling effect of 1/114) with excellent test-retest reliability (ICC = .978). Spearman rank correlations between ACTIVLIM-Hemo and HAL sub-scores ranged between .623 and .869.

Conclusions: The ACTIVLIM-Hemo is an easy-to-administer, valid and reliable alternative to HAL in assessing activity limitations in PwH. Its invariant scale can be used across conditions and time to compare the functional status of PwH over a wide measurement range.

KEYWORDS

bleeding disorder, disability and health, haemophilia, International Classification of Functioning (ICF), outcome assessment, questionnaire, rash analysis, scale

1 | INTRODUCTION

In people with haemophilia (PwH), most bleeding episodes involve the musculoskeletal system, with progressive joint destruction and functional limitations. Measuring the outcomes of haemophilia therapy requires valid and reliable assessment tools to quantify treatment benefits. With the advent of multiple novel therapies, measuring treatment effectiveness mainly based on annual bleeding rates is no longer sensitive enough.[1]

The World Federation of Haemophilia (WFH) guidelines[2] and World Health Organization's International Classification of Functioning, Disability and Health (ICF)[3] have recommended patient global health assessments that pertain to body structures and functions, such as for instance radiological scores and articular range of motion, in addition to activities like climbing stairs and participation, such as getting involved in a team sport. "Activity limitations" refers to a subject's (in)ability to execute daily activities, combining motor function and compensatory behaviours.

Activity limitations in PwH are commonly assessed with the Functional Independence Score in Haemophilia (FISH) and the Haemophilia Activities List (HAL).[4,5] The FISH is a 8-item performance-based test measuring the patient's independence in performing activities of daily living, transfers and mobility. It is calibrated to assess patients with significant musculoskeletal impairment and is therefore hardly used in patients who have access to haemophilia treatments. Its responsiveness and internal consistency are poorly documented.[6] The HAL is a self-reported questionnaire comprising 42 items covering seven domains of daily life activities. Each item is scored on a categorical scale, ranging from "Impossible" (scored 1) to "Never problems" (scored 6). Individual item scores are aggregated into three sub-scores and a sum score. Nevertheless, several methodological limitations impact the interpretation of categorical scores. Their most common drawback is their lack of unidimensionality as HAL also includes items on participation (e.g., sports and going out)[6,7] and of linearity.[8] When using the HAL, linearity problems occur when a progress from "Always problems" (scored 2) to "Mostly problems" (scored 3) in conducting an activity is interpreted the same way as a progress from "Rarely problems" (scored 5) to "Never problems" (scored 6).[8] Additionally, with the improvement of therapies and their increased availability, the activities included in existing instruments have become less challenging for PwH with a better joint status, especially for young adults who have benefitted from efficient treatments since childhood.

Modern psychometric approaches avoid such limitations by using a probabilistic model, with the Rasch model as the most established approach to objectively assess patients' perceived outcomes.[9,10] If applied to activity limitations, the Rasch model assumes that the lower the activity limitations in PwH, the higher their probability to successfully achieve any item. Using this model, an activity limitation scale can be constructed where items are located at difficulties that are invariant for all PwH, resulting in an unidimensional instrument to quantify activity limitations.[9] By selecting the items to fit a unidimensional scaling, categorical scores can be converted into a linear unit, the logit, which expresses the pass/fail probability ratio for PwH to succeed in any item. Although activity limitations can be measured as a single domain, it is also necessary to identify activities that are targeted and clinically meaningful to PwH with a better joint status.

The ACTIVLIM is a Rasch-built measure of activity limitations, previously validated in neuromuscular disorders,[11] adult stroke,[12] and child cerebral palsy.[13] Following published recommendations on the application of Rasch analysis,[14] this study sought to develop ACTIVLIM-Hemo, an ACTIVLIM scale specifically adapted to assess daily activity limitations in adult PwH.

2 | METHODS

2.1 | Participants

Overall, 120 PwH of the Belgian Haemophilia Comprehensive Treatment Centres of the Cliniques universitaires Saint-Luc Brussels and Universitair Ziekenhuis Antwerpen were included between June 2020 and December 2021. Inclusion criteria were males aged over 18 years

with mild, moderate or severe haemophilia A or B, with unmodified haemophilia therapy over the last 6 months. Two subjects reporting an haemarthrosis during the preceding month were excluded as the resulting treatment including immobilization might bias the patients' response to the self-reported questionnaires. The study was approved by local ethical committees of both hospitals involved in the study (2019/28OCT/469, B3002011942304). Written informed consent was requested to all participants. Four subjects refused to participate in the study because of lack of ability with reading documents in French or Dutch. A total of 114 patients were accepted to participate in the study and their demographic, anthropometric and clinical parameters, as well as co-medication and surgical history are presented in Table 1.

2.2 | ACTIVLIM adaptation

The ACTIVLIM-Hemo questionnaire was designed to cover the widest range of daily life activities for adult PwH. A number of 78 non-redundant activities were selected from the HAL,[5] ACTIVLIM-stroke,[12] ABILHAND,[15,16] DASH,[17] WOMAC,[18] IKDC,[19] Marx Activity Rating Scale,[20] LEFS[21] and SPADI[22] questionnaires to cover activities of self-care, dressing, in- and out-door displacements, activities of daily life and physical activities. A number of 36 additional items were devised to target activities that are clinically relevant for PwH and challenging enough for subjects with limited joint impairments. The initial questionnaire of 114 items was submitted to 19 haemophilia experts, including 5 physiotherapists, 2 haematologists and 12 PwH. They determined both the relevance and difficulty of each activity. The experimental questionnaire of 92 items was obtained after removing 35 non-relevant items, and by adding 13 additional items proposed by the panel.

2.3 | Data collection

A physiotherapist (VAC) performed the Haemophilia Joint Health Score (HJHS) 2.1 to assess joint heath[23]; the subjects completed the HAL[5] and Brief Pain Inventory short form (BPI)[24] in order to characterize the sample of PwH.

For the first assessment (T1), the experimental ACTIVLIM-Hemo was applied. Participants were invited to self-report their perceived difficulty in performing each activity without technical or human assistance, based on a three-level scale: "impossible" (scored 0), "difficult" (scored 1), or "easy" (scored 2). Activities that were unfamiliar or not performed during the last 3 months were coded as missing responses. The questionnaire was available in French and Dutch. Subjects were requested to fill in a blank form and return another form within 10 days. The latter comprised the T2 data to assess the test-retest ACTIVLIM-Hemo reliability. The questionnaire has been administered, both T1 and T2, in 1 out of 10 different random orders to avoid any systematic effect of the item presentation sequence.

TABLE 1 Characteristics of the haemophilia group ($n = 114$)

Age (years)	44 [28;60]
Weight (kg)	80 [69;90]
Height (m)	1.77 [1.73;1.81]
BMI (kg/m^2)	25.4 [22.6;29.1]
Haemophilia type	
A	89
B	25
Factor deficiency	
Severe (FVIII or IX activity < 1 IU/dl)	75
Moderate (FVIII or IX activity between 1 and 5 < IU/dl)	26
Mild (FVIII or IX activity 5 ≥IU/dl)	13
Current medical treatment	
On-demand	31
Prophylaxis	63
Emicizumab	19
Gene therapy	1
Spoken language	
French	98
Dutch	16
Lower limb articular status	
Unilateral THR	6
Bilateral THR	1
Unilateral TKR	16
Bilateral TKR	13
Unilateral ankle arthrodesis	8
Bilateral ankle arthrodesis	1
Unilateral TAR	2
Haemophilia Joint Health Score 2.1	
Elbow score (/40 pts)	3 [1;14]
Knee score (/40 pts)	3 [1;12]
Ankle score (/40 pts)	7 [1;16]
Knee + ankle joint score (/80 pts)	11 [4;26]
Total articular score (/120 pts)	15 [5;41]
Haemophilia Activities List ($n = 110/114$)	
Upper extremity activities score (/100 pts)	86 [69;96]
Basic lower extremity activities score (/100 pts)	70 [47;92]
Complex lower extremity activities score (/100 pts)	60 [29;86]
Total score (/100 pts)	76 [54;90]
Haemophilia activities list short form ($n = 110/114$) (/100pts)	76 [59;92]
Brief PAIN INVENTory severity score ($n = 110/114$) (/10 pts)	1.8 [.4;3.4]

Values are *N* or median [P25;P75].
Abbreviations: BMI, Body Mass Index; TAR, total ankle replacement; THR, total hip replacement; TKR, total knee replacement.

2.4 | Data analysis

2.4.1 | Measuring activity limitations with the Rasch Rating Scale model

The dataset was analysed using RUMM2030 (RUMM Laboratory Pty Ltd, Perth, Western Australia) with the Rasch Rating Scale model.[11,26,27] When modelling the activity's difficulty as perceived by a subject, the observed response solely depended on[1] subject's activity limitation[2]; difficulty of activity[3]; threshold location between proposed responses (two thresholds per item: one between "impossible" and "difficult"; one between "difficult" and "easy"). The three parameter types were determined by RUMM to locate each subject, activity and threshold on a common measurement scale. Thresholds were located where the probabilities of choosing one of both adjacent responses to an item were equal. At the first threshold, PwH displayed an equal probability of choosing "impossible" and "difficult" for that item. At the second threshold, they displayed an equal probability of choosing "difficult" and "easy" for that item. Each item was located at the average location of both thresholds. The location range of all item thresholds defined the measurement range of the resulting scale. The measurement scale was calibrated in logits, meaning a linear probabilistic unit expressing the pass/fail probability ratio for a subject to succeed any item.[27]

2.4.2 | Verifying the scale's psychometric properties

Based on the Rasch model, the scale's psychometric properties could be verified, with each item contributing to defining a targeted, unidimensional and invariant scale of activity limitations.[26] After determining the locations of PwH, items and thresholds, the model could calculate the expected response for each subject to each item. The difference between the observed and expected score was used to calculate a standard residual statistic.[25,28] This revealed whether the response to an item fitted the model's expectation, with each item fitting the unidimensional scale being defined by the others. If not, over-discriminating items displayed more predictable than expected responses, which could, however, be kept if sufficiently informative. Under-discriminating items constituted a higher threat to the scale's unidimensionality, and were thus deleted. To assess any misfit's significance, Chi-squared statistics were applied, with significance levels set at .05.

2.4.3 | Selecting activities that define ACTIVLIM-Hemo

Following successive analyses, RUMM software was used to select the activities retained to delineate a targeted unidimensional and unbiased scale of activity limitations in PwH.[14] Selected items had to[1] be targeted to PwH activity levels[2]; fit the unidimensional scale defined by all other selected items and[3] verify the assumption of local independence by being not linked more than by their difficulty (residual correlations <.4).[14] The absence of differential item functioning was checked graphically[29] by observing a stable item difficulty hierarchy across different demographic and clinical subject subgroups dichotomously split according to the following factors: age (median: 44 years), body mass index (median: 25.4), pain (median BPI score:1.75/10), lower limb arthropathy (≥one lower limb joint with HJHS above the median 2), lower limb surgery (≥one lower limb surgery) and time (T1 vs. T2).

2.4.4 | Interpreting ACTIVLIM-Hemo measures of activity limitations

As with any Rasch analysis, the resulting scale origin of 0 logit was conventionally set at the average item difficulty. Subjects' activity levels with negative logit locations corresponded to less able PwH, positive logit locations corresponded to more able PwH. Items with a higher logit location are more difficult items, lower logit locations correspond to easier items. Given the scale's linear nature, at any activity level, a one logit difference between two subjects indicated that their odds of successful activity achievement were in a ratio of 2.71 (e^1). For interpretation convenience, the logit scale was converted into a measurement range percentage, where 0% corresponds to the smallest measurable level of activity and 100% corresponds to the highest measurable level of activity, with average item difficulty approximating 50%.

2.4.5 | Reliability

The ACTIVLIM-Hemo reliability in PwH was computed using the person separation index (PSI), indicating the ratio of true measure versus observed variance corrected for measurement error.[26] This index can be determined in the presence of missing responses and accounts for a measurement error computed specifically for each subject, allowing the computation of the number of activity level groups that are statistically distinguished within the sample.[30] Hence, the test-retest reliability could be approached in three ways. First, an individual approach would compare the subject's measures at T1 and T2 relative to their measurement error. Second, the intraclass correlation coefficient (ICC) between single measurements was computed using a two-way mixed effect model for absolute agreement between T1 and T2 data.[31,32] Third, the equivalence of PwH activity level at T1 and T2 was assessed graphically by Bland and Altman plotting.[33]

2.4.6 | Construct validity

Convergent validity was assessed with Spearman rank correlation coefficients between the activity limitation measured by

TABLE 2 ACTIVLIM-Hemo calibration

Item	Label	Difficulty (%range)	SE	Difficulty (logits)	SE	Fit residual	Chi square	dof[a]	*p*-value
a.	Skiing	76.3	3.7	3.107	.471	−.444	2.491	2	.288
b.	Sprinting	76.3	2.6	3.105	.329	−.788	2.922	2	.232
c.	Jogging	71.0	1.9	2.440	.241	−1.343	6.981	2	.030
d.	Running down a steep slope	67.9	2.5	2.043	.315	−.678	1.364	2	.506
e.	Running in the woods	67.7	1.9	2.029	.242	−.642	2.184	2	.336
f.	Running	67.4	1.8	1.984	.222	−.488	4.768	2	.092
g.	Jumping as high as possible	67.2	2.5	1.955	.318	−.855	1.090	2	.580
h.	Jumping down from a chair	64.4	1.9	1.608	.233	−.717	3.028	2	.220
i.	Hopping on one foot	61.8	1.8	1.281	.224	−.559	1.339	2	.512
j.	Playing a racket sport (e.g. tennis, table tennis, badminton, squash)	57.9	2.0	.786	.248	2.723	1.493	2	.474
k.	Doing a sport more than once a week	54.4	1.9	.344	.235	1.066	2.940	2	.230
l.	Bicycling in the woods	52.7	1.8	.129	.231	.831	2.533	2	.282
m.	Standing on tiptoes	50.2	1.7	−.185	.213	2.221	2.185	2	.335
n.	Kicking a soccer ball	45.1	1.8	−.824	.231	−.540	.336	2	.845
o.	Climbing a ladder	43.3	1.8	−1.044	.228	1.414	1.183	2	.553
p.	Climbing stairs	40.3	1.8	−1.421	.230	.369	3.567	2	.168
q.	Standing on a chair	40.3	1.9	−1.421	.233	−1.003	1.336	2	.513
r.	Replacing a ceiling light bulb	32.5	2.8	−2.403	.353	−.348	.099	2	.952
s.	Walking outside on flat ground	27.3	2.3	−3.065	.295	.246	1.859	2	.395
t.	Buttoning up a shirt	26.8	2.4	−3.126	.299	.066	.984	2	.611
u.	Walking 100 meters	23.5	2.6	−3.537	.328	−.115	.646	2	.724
v.	Putting on a pair of trousers	21.5	2.8	−3.787	.349	−.299	.203	2	.903
Mean		51.6	2.2	.000	.276	.005			
Standard deviation		17.7	.5	2.224	.064	1.051			

The item fit has been assessed over 3 CI in order to balance the number of subjects in each CI (ca. 38 subjects per CI).
[a]Degrees of freedom are the number of class intervals (CI) −1.

ACTIVLIM-Hemo at T1 and HAL sub-scores and HAL short form.[34] Spearman's correlations were preferred over Pearson's because the HAL scores are ordinal in nature and do not meet the requirements for linear correlations. We hypothesized that at least a moderate significant correlation would be observed since the HAL is considered as the reference patient-reported questionnaire for activity limitation in PwH.[6] Discriminative validity was assessed by testing the relationship between ACTIVLIM-Hemo and demographic and clinical indicators using independent *t*-tests or univariate ANOVA with Bonferroni post-hoc pairwise comparisons, with statistical significance set at .05. We hypothesized that a significant difference would be observed as a function of age,[35,36] factor deficiency[35] and the extent of joint damage[36] but not between types of haemophilia.[37] Although we expect a higher activity limitation in obese PwH (BMI > 30), we only tested the relation of activity limitation to BMI as a description of the clinical profile in our sample.

3 | RESULTS

3.1 | Refinement of ACTIVLIM-Hemo scale for PwH

Following successive Rasch model-based analyses, 22 items were selected, delineating a targeted, unidimensional and invariant scale of activity limitations in PwH. From the 92-item experimental ACTIVLIM-Hemo questionnaire, 21 items were deemed mis-targeted and too easy (e.g. "standing up from sitting on a chair"), 36 were not part of the construct defined by the other items (e.g. "tying shoe laces"), 2 were over-discriminating with highly predictable responses (e.g. "walking more than 1 km"), while 11 items were deleted due to being more linked with another item than by their difficulty (e.g. "practicing a sport once a week" was deleted in favour of "practicing a sport more than once a week").

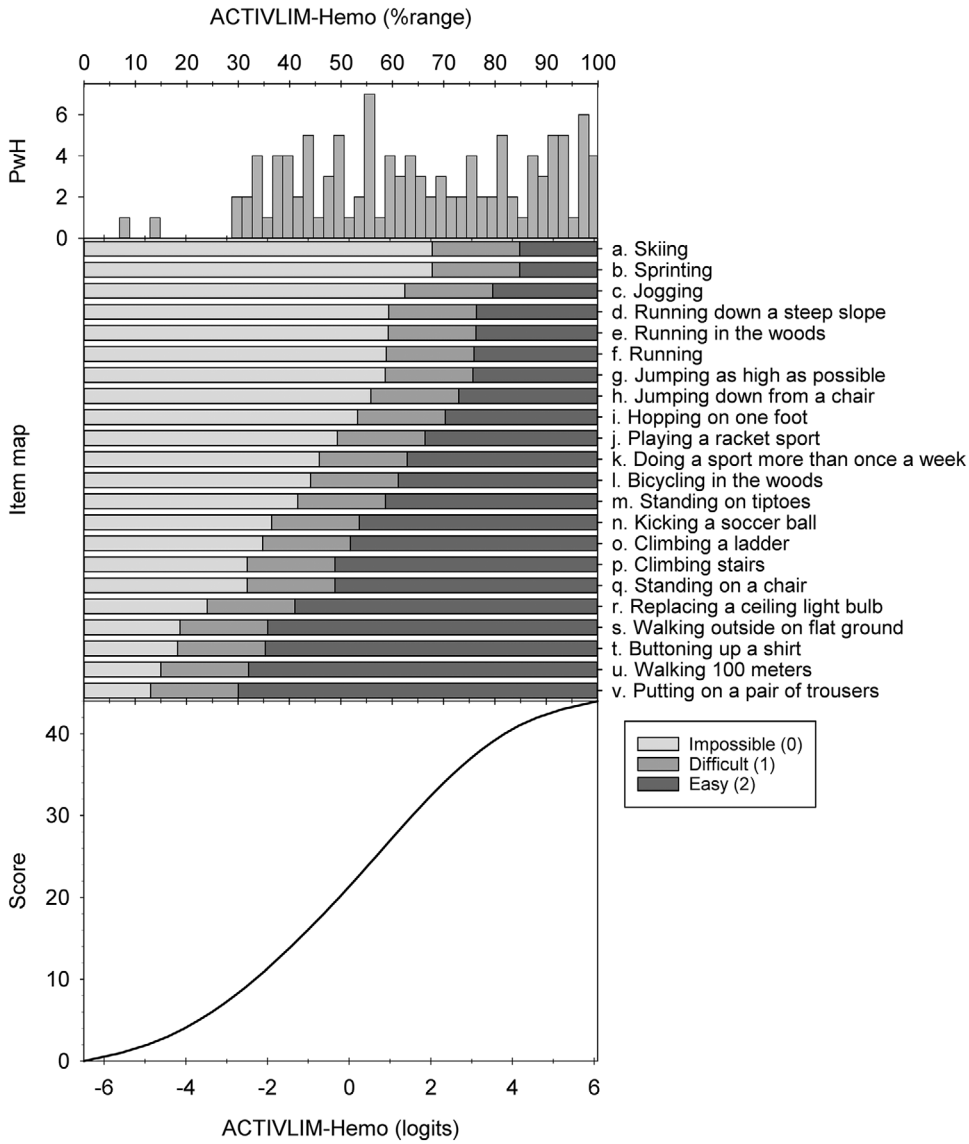

FIGURE 1 The activity limitation measured using ACTIVLIM-Hemo is presented on the abscissa graduated in % range (scale measurement range, top axis) and in logits (the pass/fail probability ratio for a subject to succeed an item, bottom axis). For both units, higher values indicate a higher activity level (right side of the figure). The distribution of activity levels in our PwH sample is shown on the top panel, where most PwH lie above an activity level of 25%, excepting two subjects with a very severe activity limitation. The difficulty hierarchy of items is shown on the item map (middle panel) listed, from top to bottom, in decreasing difficulty order. The item map shows the activity level required to succeed each item, as follows. For any item, a PwH with a severe activity limitation is expected to answer "impossible" (light grey, score of 0); a PwH with a moderate activity limitation is expected to answer "difficult" (medium grey, score of 1); a PwH with a mild or no activity limitation is expected to answer "easy" (dark grey, score of 2). The item map displays the thresholds that separate consecutive response categories, being located from −8.5 to +8.5% range (−1.074 to +1.074 logits) in regard to the difficulty of each item. The item map also shows that the distance between thresholds is similar across activities, lying within 2 SD of the mean distance for all items. The conversion of raw ordinal scores to linear measures is shown on the bottom panel across the whole measurement range (from a score of 0 indicating that all items are "impossible" to a score of 44, i.e., 22*2, indicating that all items are "easy"). Given the S-shaped relationship, a gain of 1 score point corresponds to different activity level changes across the measurement range. At the centre of the measurement, a change score from 21 to 22 results in difference of 1.4% range (.18 logits) in activity limitation, while at the measurement range's extremity, the same change score, from 1 to 2, results in an over three-fold difference in activity of 5.2% range (.66 logits), highlighting the non-linearity of the scores.

3.2 | Psychometric properties of ACTIVLIM-Hemo

Table 2 presents the calibration of the final 22-item ACTIVLIM-Hemo, with items listed in order of decreasing difficulty. The item difficulty and standard error (SE) were listed in logits and % of measurement range, higher values representing more difficult items. Both units were constant along the measurement range, with the scale centred on the average item difficulty conventionally set at 0 logits, and % range expanding from 0% to 100% of measurement range. In Table 2, standard residual and chi-squared fit statistics are presented, with

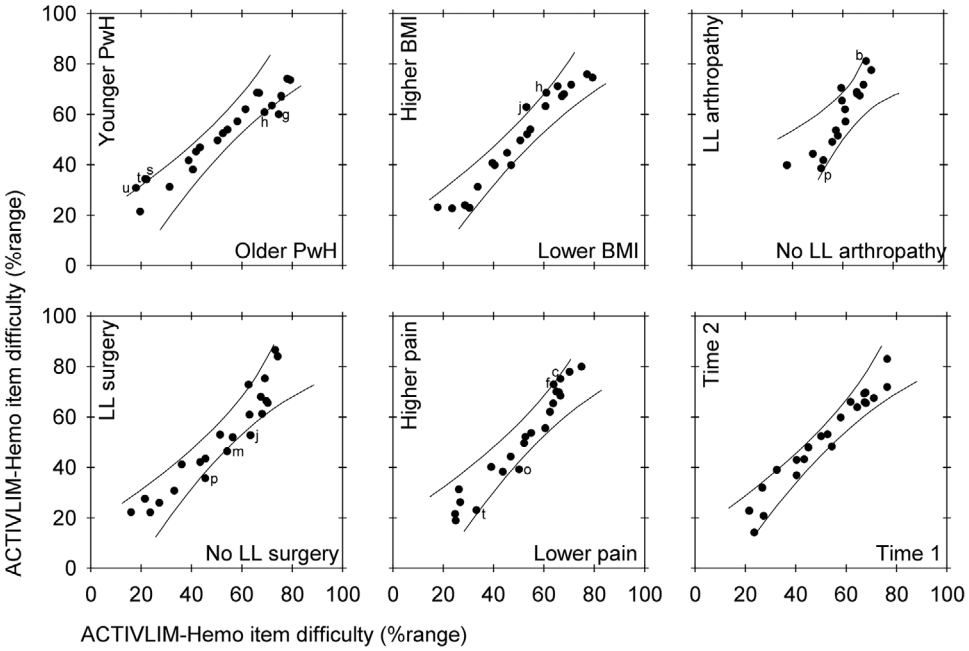

FIGURE 2 Differential Item Functioning (DIF) plots showing the difficulty of the 22 ACTIVILIM-HEMO items when estimated on sub-samples split according to sex different criteria: age split on the median of 44 years; body mass index (BMI) split on the median of 25.4; lower limb (LL) arthropathy (split on the presence of at least one LL joint with a HJHS above the median score of 2), LL surgery split on the presence of at least one surgery at the LL; pain split on the median BPI score of 1.8; time split on T1 versus T2. In each plot, more difficult items are plotted in the top/right part of the graph. Most items lie within the 95% confidence intervals (CI, solid lines) of the ideal invariance, indicating that the ACTIVLIM-Hemo provides an invariant measure of activity limitation unbiased by either demographic or clinical conditions. Outliers are identified by their label.

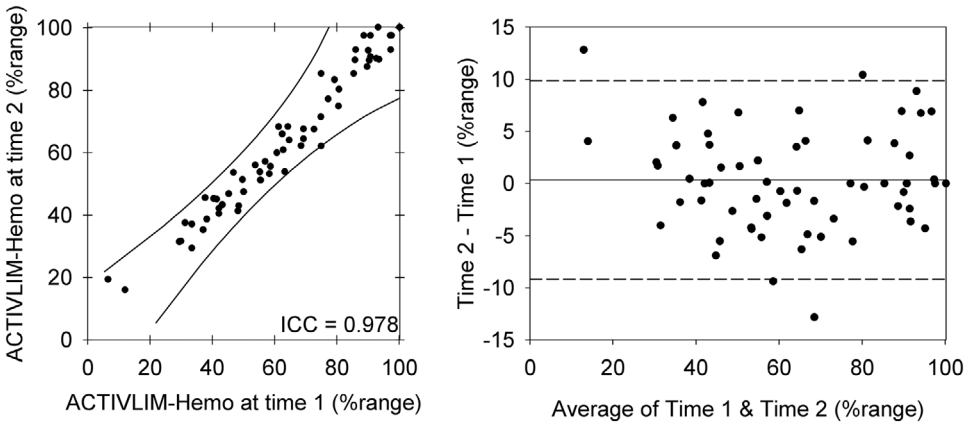

FIGURE 3 Test-retest reliability illustrated via the individual approach (left panel) following a median delay of 12 days. The activity level for each PwH (dots) is compared between T1 and T2, where higher activity levels are plotted in the top-right part of the graph. All measures lie within the 95% confidence interval (CI) of the ideal equivalence over time. The intraclass correlation coefficient (ICC = .978) also confirms an excellent test-retest reliability at group level. The Bland and Altman plot (right panel) shows an overall agreement at both times based on an average difference of .35% range (plain line), while most differences lie within the 95% confidence interval (CI, dashed lines) placed at 1.96 times, the SD (4.86% range) being around the mean difference.

associated *p*-values. All 22 items define a unidimensional scale of activity limitations in PwH, displaying the expected item discrimination. Only item c, "jogging," displayed a discrimination significantly higher than expected, yet its limited amplitude did not threaten the scale's unidimensionality. The final PSI was equal to .92. Hence, 4–5 activity level groups could be statistically distinguished within the sample.

3.3 | ACTIVLIM-Hemo description

The ACTIVLIM-Hemo is illustrated in Figure 1, where PwH exhibited a median [IQR] response rate of 95% [87%; 99%] to all items. The PwH activity levels ranged from 7.1% to 100% range (−5.60 to 6.09 logits), overall activity limitation being moderate with 64.7% range on

average (1.65 logits). A minor ceiling effect was observed in one subject reporting all activities as "easy." The selected items were thus adequately targeted to PwH's activity levels. The item map highlighted the disablement pattern in PwH, "skiing" being the most challenging activity (76.3% range) and "putting on a pair of trousers" the easiest one (21.5% range). The most probable score to each item was computed as a function of the subject's activity limitations (Figure 1). Thus, "walking outside on flat ground" easily corresponded to an activity level of about 35% range, whereas "standing on tiptoes" easily reflected a higher activity level of about 60% range. Whatever the subject's limitations, "standing on tiptoes" was about 20 (e^3) times more difficult than "walking outside on flat ground" in PwH. Figure 1 illustrates the S-shaped relationship used to convert ordinal raw scores into linear measures of activity limitation.

To test whether the difficulty hierarchy of the 22 retained activities was similar across PwH subgroups, we defined dichotomous PwH categories according to six splitting criteria. The perceived difficulty of each item contrasted between both mutually exclusive categories in the differential item functioning plots of Figure 2. Since most items lied within the 95% confidence interval (CI) of ideal identity, the perceptions were likely group independent, with the difficulty hierarchy of ACTIVLIM-Hemo uniformly perceived by PwH, yet with a few exceptions. Indeed, "jumping as high as possible" was relatively more difficult for older PwH, whereas "climbing stairs" was relatively easier for PwH with prior lower limb surgery.

3.4 | Test-retest ACTIVLIM-Hemo reliability

Overall, 63 PwH returned their re-assessment form following a median [IQR] of 12 [7;27] days, with ACTIVLIM-Hemo test-retest reliability illustrated in Figure 3. Using the individual approach, all activity limitation measures lied within the 95% CIs of ideal equivalence over time. Accordingly, the .978 ICC was indicative of an excellent test-retest reliability at group level. The overall agreement between measures at both times was confirmed by Bland and Altman plotting, with an average .35% difference range over time, most differences lying within 95% CIs.

3.5 | Construct validity of ACTIVLIM-Hemo

Correlations between ACTIVLIM–Hemo measures and HAL sum score, HAL sub-scores and HAL short form are illustrated in Figure 4. All correlations were highly significant ($p < .001$), with high to moderate amplitude. Higher correlations ($\rho > .8$) were observed with HAL sum score, lower limb complex score and short form, and lower correlations ($\rho = .623$) with the upper limb sub-score. The ceiling effect observed in HAL was similarly highlighted, being 9% in HAL sum score, 17% in upper limb and lower limb basic sub-scores, 12% in lower limb complex sub-score and 7% in short form. Only one PwH (1%) displayed a ceiling effect in ACTIVLIM-Hemo.

The impact of demographic and clinical variables on ACTIVLIM-Hemo measures is presented in Table 3. A significant difference in ACTIVLIM-Hemo measure was observed according to age, factor deficiency, and HJHS. No significant differences were noted between both haemophilia types, with trends toward differences depending on BMI.

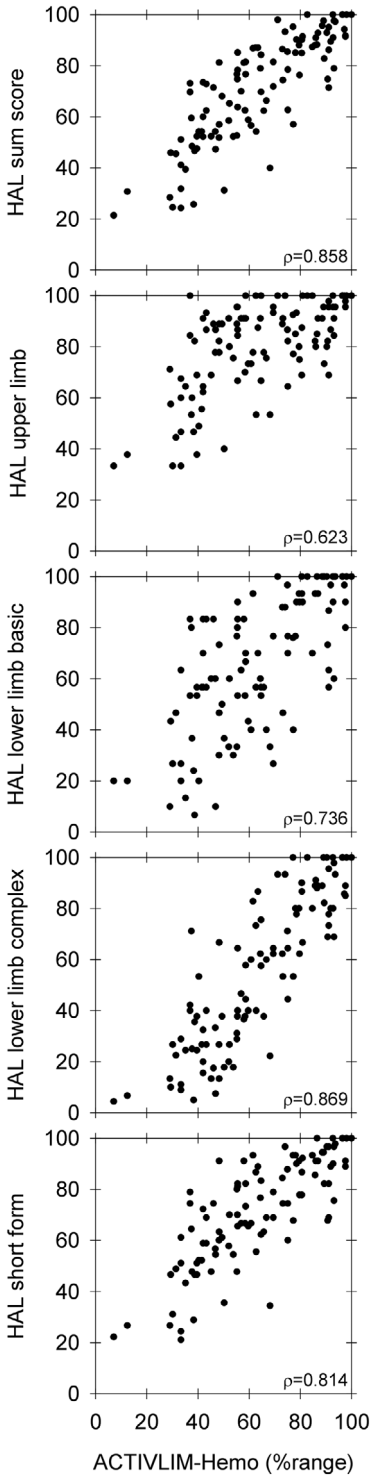

FIGURE 4 Correlation between the linear measure of ACTIVLIM-Hemo (abscissa), ordinal HAL sum score and sub-scores and HAL short form (ordinates) for each PwH (dots). Spearman's rank order correlation coefficients are presented for each HAL sub-score and for the short form.

TABLE 3 Relationship of ACTIVLIM-Hemo to demographic and clinical indices

		ACTIVLIM-Hemo (%range)								Effect	
Variable	N	Mean (SD)	Posthoc sign.							Statistic	*p*-value
Age											
Younger than 44 years	57	74.9 (18.8)	N/A							$t = 5.436$	<.001
44 years or more	57	54.6 (21.0)									
BMI											
Lower than 25.4	57	68.4 (23.0)	N/A							$t = 1.768$	.080
25.4 or more	57	61.1 (21.1)									
Haemophilia type											
A	89	64.1 (22.5)	N/A							$t = .584$	.560
B	25	67.0 (21.9)									
Factor deficiency											
Severe	75	60.4 (23.1)				*				$F = 4.376$	.015
Moderate	26	73.6 (18.9)				*					
Mild	13	71.9 (17.6)									
Joint damage											
All joints HJHS ≤ 2	26	85.5 (10.9)	*	*						$F = 32.163$	<.001
At least 1 UL joint with HJHS > 3	9	88.6 (10.8)				*	*				
At least 1 LL joint with HJHS > 3	21	64.8 (17.0)	*			*			*		
At least 1 UL and 1 LL joint with HJHS > 3	58	51.7 (18.9)		*			*		*		

Abbreviations: BMI, Body Mass Index; LL, Lower Limb; N/A: not applicable; UL, Upper Limb.*Indicates pairwise Bonferroni post-hoc significant differences.

4 | DISCUSSION

The ACTIVLIM questionnaire was adapted for assessing activity limitations in PwH. Through selection of activities based on existing tests and addition of activities deemed relevant to PwH, the ACTIVLIM-Hemo was built using the Rash model. The rationale for developing a new measure of activity limitations specific to PwH was to determine the difficulty perceived for each activity in this clinical condition. Another motivation was to create a new scale that offered more challenging activities than the exiting instruments, such as HAL[5] and FISH,[4] in order to better differentiate PwH with rather good joint conditions. The final 22-item questionnaire delineated a single dimension, covering the whole spectrum of pertinent activities, where PwH can be located from more to less limited on a linear scale. The ACTIVLIM-Hemo questionnaire is extremely easy to administer, as it can be self-completed within five minutes, with scores being readily analysable on www.rehab-scales.org as illustrated in Figure 5.

The ACTIVLIM-Hemo results were shown to reflect a subject's perceived difficulty to perform daily activities. Unlike functional tests, the ACTIVLIM-Hemo provided a measure of the subjects' perceptions within their own environment, unbiased by either instrumental setup or evaluator, yet an ecological functional measure of PwH in their daily life.[38] The difficulty of ACTIVLIM-Hemo activities were likely perceived invariantly across PwH subsamples differing in clinical presentations. The observed invariance over time supported an excellent test–retest reliability (ICC = .978). The measurement range of final ACTIVILIM-Hemo was wide enough to distinguish clinically relevant functional limitations, with levels including challenging activities like "sprinting," being thus able to differentiate PwH with lower limitations; likewise, easier activities like "walking 100 meters" allowed for differentiating PwH with higher limitations. Its internal consistency was high (PSI = .92), with only a minor ceiling effect (1/114 PwH).

The ACTIVLIM-Hemo delineates the challenge imposed on PwH by each activity, highlighting the typical disablement pattern in this population. Activities mainly involving the upper limbs can be conducted by primarily using the less impaired limb in a compensation strategy, thus being reported as the easier ones. Activities involving the weight-bearing capacity require both lower limbs and were reported as the most challenging ones. Importantly, the hierarchy of activity difficulties constitutes a unidimensional gradient, where PwH who progress upon easier activities do also increase their probability to succeed more difficult ones.

The relationship between clinical findings and activity limitations does not only constitute a validation of ACTIVLIM-Hemo, but it is a clinical end point in itself. The activity limitations measured herein were in line with previous reports indicating that older PwH with no previous access to replacement therapy were more disabled than younger ones,[35,36] while severe PwH displayed a higher activity limitation than moderate ones.[35] Similarly, arthropathy extent was shown to correlate with activity limitations,[36] PwH with arthropathies at both upper and lower limbs displaying higher activity limitations than those

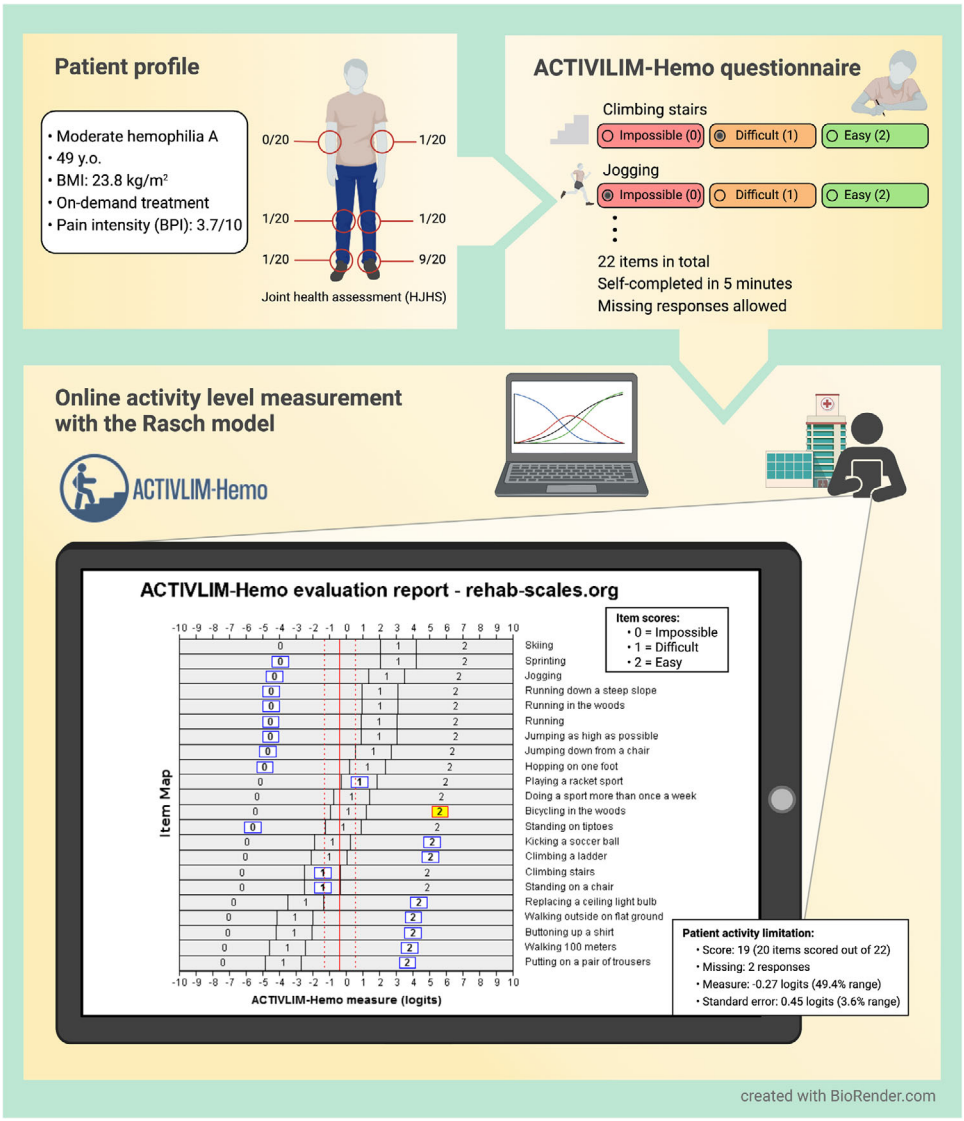

FIGURE 5 A sample scoring form showing the evaluation of a subject with haemophilia. The online scoring form is presented to the subject in random orders to avoid any systematic effect of the item presentation sequence. The score interpretation (bottom image) shows items ordered from most (top) to least (bottom) difficult, where possible responses to each item are identified by the corresponding score (0 for "impossible," 1 for "difficult" and 2 for "easy") as a function of the subject's activity limitation expressed in logits. The red vertical line indicates the subject's activity limitation of −.27 logits (49.4 %range); the dashed vertical red lines indicate its 95% confidence interval between −1.16 and .61 logits (42.4–56.5 %range). The figure confirms that, given the subject's activity limitation, the expected response is observed for most items of the ACTIVLIM-Hemo questionnaire (responses in blue squares), hence confirming the coherence of the subject's responses. The response to the item "Bicycling in the woods" is slightly unexpected (responses in yellow square) since the subject reports it as "easy" while it should be "difficult" or "impossible" given the 95% confidence interval of the subject's activity limitation. Note that the items "Skiing" and "Doing a sport more than once a week" are not answered, yet not compromising the interpretation of the subject's activity limitation.

with only lower limb involvement. A trend towards a higher activity limitation in PwH with higher BMI was similarly observed.[39]

The ATIVLIM-Hemo validity is further reinforced by its correlation with HAL, ACTIVLIM-Hemo sharing 74% of variance (ρ^2) in common with its sum score, and 66% with the HAL short form score. ACTIVLIM-Hemo's ceiling effect (<1%) turned out to be lower than that of HAL (9%), especially concerning the less challenging upper limb sub-score (17%). With ACTIVLIM-Hemo gathering all information on a subject's activity limitation within a single measure, its measurement range extended beyond that of HAL, attaining higher activity levels. The highest correlation was observed with the HAL lower limb complex sub-score, supporting the scale's application in PwH treated with novel therapies who will likely report minor activity limitations.

Besides its wider measurement range, ACTIVLIM-Hemo offers additional advantages over other self-reported questionnaires. Its major asset is the conversion of the ordinal raw scores into a true linear unit than can be expressed in % measurement range, enabling quantitative comparisons between different treatments and over time.[8,27]

ACTIVLIM-Hemo contains a three-level response scale, offering sufficient information in a reliable self-reported format with response categories that are easily distinguished.[40] On account of its tolerance for missing responses, ACTIVLIM-Hemo allows for calculating subjects' activity limitation even if they omit filling in unusual activities.

5 | CONCLUSION

ACTIVLIM-Hemo is a valid alternative to HAL for assessing activity limitations in PwH. It offers an invariant scale that can be used across demographic and clinical conditions to compare the functional status of PwH over a wide measurement range, with sufficient reliability to assess treatment responses in future studies, this being only one of its several future applications.

AUTHOR CONTRIBUTIONS

SL and MP designed the research study, analysed the data and wrote the paper. AF and VAC acquired the data. All co-authors reviewed the manuscript.

ACKNOWLEDGEMENTS

This study was supported by the ASPIRE grant from Pfizer Inc. (grant number 55563735). The authors are grateful to Norman Heglund for reviewing the English translation of the questionnaire and Evelien Krumb for the realisation of the figures.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare. All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report. We certify that the submission is original work and is not under review at any other publication.

ETHICS STATEMENT

The study was approved by local ethical committees of both hospitals involved in the study (2019/28OCT/469, B3002011942304).

ORCID

Sébastien Lobet https://orcid.org/0000-0002-3829-6850
Catherine Lambert https://orcid.org/0000-0003-2222-0357
Anthe Foubert https://orcid.org/0000-0002-2514-9026
Valérie-Anne Chantrain https://orcid.org/0000-0002-2044-7112
Nathalie Roussel https://orcid.org/0000-0002-7409-878X
Mira Meeus https://orcid.org/0000-0002-2022-5114
Cedric Hermans https://orcid.org/0000-0001-5429-8437
Massimo Penta https://orcid.org/0000-0002-0355-8638

REFERENCES

1. Konkle BA, Skinner M, Iorio A. Hemophilia trials in the twenty-first century: defining patient important outcomes. *Res Pract Thromb Haemost [Internet]*. 2019;3(2):184-192. [cited 2022 May 16].
2. Srivastava A, Santagostino E, Dougall A, et al. WFH guidelines for the management of hemophilia. *Haemophilia*. 2020 ;26(S6):1-158.
3. WHO. The international classification of functioning, disability and health. *World Heal Organ*. 2001;18:237.
4. Poonnoose PM, Manigandan C, Thomas R, et al. Functional independence score in haemophilia: a new performance-based instrument to measure disability. *Haemophilia*. 2005;11(6):598-602.
5. van Genderen FR, van Meeteren NLU, van der Bom JG, et al. Functional consequences of haemophilia in adults: the development of the haemophilia activities list. *Haemophilia*. 2004;10(5):565-571.
6. Timmer MA, Gouw SC, Feldman BM, et al. Measuring activities and participation in persons with haemophilia: a systematic review of commonly used instruments. *Haemophilia [Internet]*. 2018;24(2):e33-e49. [cited 2022 Oct 22].
7. Poonnoose P, Keshava S, Gibikote S, Feldman BM. Outcome assessment and limitations. *Haemophilia*. 2012;18(SUPPL.4):125-130.
8. Merbitz C, Morris J, Grip J. Ordinal scales and foundations of misinference. *Arch Phys Med Rehabil*. 1989;70(4):308-312.
9. Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests*. MESA Press; 1980.
10. Andrich D. A rating formulation for ordered response categories. *Psychometrika*. 1978;43(4):561-573.
11. Vandervelde L, Van den Bergh PYK, Goemans N, Thonnard JL. ACTIVLIM: a Rasch-built measure of activity limitations in children and adults with neuromuscular disorders. *Neuromuscul Disord*. 2007;17(6):459-469.
12. Batcho CS, Tennant A, Thonnard JL, ACTIVLIM-Stroke: a crosscultural Rasch-built scale of activity limitations in patients with stroke. *Stroke [Internet]*. 2012 ;43(3):815-823. [cited 2022 Feb 23]
13. Bleyenheuft Y, Paradis J, Renders A, Thonnard JL, Arnould C. ACTIVLIM-CP a new Rasch-built measure of global activity performance for children with cerebral palsy. *Res Dev Disabil [Internet]*. 2017;60:285-294. [cited 2022 Feb 23].
14. Tennant A, Conaghan PG. The rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care Res*. 2007;57(8):1358-1362.
15. El Khoury G, Penta M, Barbier O. ABILHAND-HS: a linear scale for outcome measurement in hand surgery. *J Hand Surg Eur Vol [Internet]*. 2021;46(6):676-678. [cited 2022 Feb 23].
16. Durez P, Fraselle V, Houssiau F, Thonnard J-L, Nielens H, Penta M. Validation of the ABILHAND questionnaire as a measure of manual ability in patients with rheumatoid arthritis. *Ann Rheum Dis*. 2007;66(8):1098-1105.
17. Hudak P, Amadio P, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). *Am J Ind Med*. 1996;29(6):602-608.
18. Bellamy N, Buchanan W, Goldsmith C, Campbell J, Stitt L. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol*. 1988;15(12):1833-1840.
19. Irrgang JJ, Anderson AF, Boland AL, et al. Development and validation of the international knee documentation committee subjective knee form. *Am J Sports Med [Internet]*. 2001;29(5):600-613. [cited 2022 Feb 23].
20. Marx RG, Stump TJ, Jones EC, Wickiewicz TL, Warren RF. Development and evaluation of an activity rating scale for disorders of the knee. *Am J Sports Med [Internet]*. 2001;29(2):213-218. [cited 2022 Feb 23].
21. Mehta SP, Fulton A, Quach C, Thistle M, Toledo C, Evans NA. Measurement properties of the lower extremity functional scale: a systematic review. *J Orthop Sports Phys Ther [Internet]*. 2016;46(3):200-216. [cited 2022 Feb 23].
22. Roach K, Budiman-Mak E, Songsiridej N, Lertratanakul Y. Development of a shoulder pain and disability index. *Arthritis Care Res*. 1991;4(4):143-149.
23. Hilliard P, Funk S, Zourikins N, et al. Hemophilia joint health score reliability study. *Haemophilia*. 2006;12(5):518-525.

24. Stanhope J. Brief pain inventory review. *Occup Med (Lond) [Internet]*. 2016;66(6):496-497. [cited 2022 Feb 28].
25. Andrich D. *Rasch Models for Measurement*. SAGE Publications; 1988.
26. Wright B, Masters G. *Rating Scale Analysis*. Mesa Press; 1982.
27. Wright B, Linacre J. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil*. 1989;70(12):857-860.
28. Andrich D, Sheridan B. *Interpreting RUMM2020: Part II, Polytomous Data*. RUMM laboratory Pty Ltd; 2005.
29. Wright B, Stone M. *Best Test Design*. Mesa Press; 1979.
30. Fisher W. *Reliability Statistics*. MESA Press; 1992;6(3):238.
31. Shrout PE, Fleiss JL:420-428.
32. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med [Internet]*. 2016;15(2):155-163. [cited 2022 Mar 15].
33. Altman D, Bland J. Measurement in medicine: the analysis of method comparison studies. *Statistician*. 1983;32:307-317.
34. Kuijlaars IAR, van der Net J, Buckner TW, Kempton CL, Schutgens REG, Fischer K. Shortening the haemophilia activities list (HAL) from 42 items to 18 items. *Haemophilia [Internet]*. 2021;27(6):1062-1070. [cited 2022 Feb 15].
35. Kempton CL, Buckner TW, Fridman M, Iyer NN, Cooper DL. Factors associated with pain severity, pain interference, and perception of functional abilities independent of joint status in US adults with hemophilia: multivariable analysis of the pain, functional impairment, and quality of life (P-FiQ) study. *Eur J Haematol [Internet]*. 2018;100(Suppl 1):25-33. [cited 2022 Mar 29].
36. McLaughlin P, Morris R, Chowdary P. Investigating the relationship between the HJHS and HAL in routine clinical practice: a retrospective review. *Haemophilia [Internet]*. 2018;24(6):988-994. [cited 2022 Mar 29].
37. Castaman G, Matino D. Hemophilia A and B: molecular and clinical similarities and differences. *Haematologica [Internet]*. 2019;104(9):1702. [cited 2022 Nov 13].
38. Grimby G, Andrén E, Daving Y, Wright B. Dependence and perceived difficulty in daily activities in community-living stroke survivors 2 years after stroke: a study of instrumental structures. *Stroke*. 1998;29(9):1843-1849.
39. Biere-Rafi S, Haak BW, Peters M, Gerdes VEA, Büller HR, Kamphuisen PW. The impairment in daily life of obese haemophiliacs. *Haemophilia*. 2011;17(2):204-208.
40. Linacre J. Investigating rating scale category utility. *J Outcome Meas*. 1999;3(2):103-122.

How to cite this article: Lobet S, Lambert C, Foubert A, et al. ACTIVLIM-Hemo: A new self-reported, unidimensional and linear measure of activity limitations in persons with haemophilia. *Haemophilia*. 2022;1-12. https://doi.org/10.1111/hae.14705