EXTENDED REPORT

Validation of the ABILHAND questionnaire as a measure of manual ability in patients with rheumatoid arthritis

Patrick Durez, Virginie Fraselle, Frédéric Houssiau, Jean-Louis Thonnard, Henri Nielens, Massimo Penta

....

See end of article for authors' affiliations
....

Correspondence to:
Patrick Durez, MD,
Department of
Rheumatology, UCL 5390,
Av Hippocrate, 10, B-1200
Bruxelles, Belgium; patrick.
durez@ruma.ucl.ac.be

Objective: Hand and upper limb involvement is common in patients with rheumatoid arthritis (RA). However, its impact on manual activities of daily life has not been fully evaluated. A measure of manual ability was developed, through the Rasch measurement model, by adapting and validating the ABILHAND questionnaire, which measures the patient's perceived difficulty in performing everyday manual activities.

Methods: 112 patients with RA were evaluated. The following tests were performed: the ABILHAND questionnaire, the Health Assessment Questionnaire (HAQ), the Jamar grip and key pinch strength tests, the Box and Block dexterity test and the Purdue pegboard dexterity test. In total, 35 patients were reassessed to determine the test–retest reliability of the ABILHAND, and 6 patients were studied before and after therapy with tumour necrosis factor (TNF) blockers to address sensitivity to change.

Results: The Rasch refinement of the ABILHAND led to a selection of 27 items rated on a 3-point scale. The resulting ability scale was targeted to the ability of the patients. The item-difficulty hierarchy was stable across demographic and clinical subgroups and over time. Grip and key pinch strength and manual and digital dexterity on both hands were significantly, though moderately, correlated with the ABILHAND measures. Manual ability was also significantly related to the number of affected hands, disease duration, tender and swollen joint counts on upper limbs, disease activity and the HAQ. Sensitivity to change was demonstrated in patients treated with TNF blockers, commensurate with their clinical improvement.

Conclusion: The ABILHAND questionnaire is a clinically valid person-centred measure of manual ability that could be useful in longitudinal RA studies.

Rheumatoid arthritis (RA) is a chronic inflammatory disease affecting multiple peripheral joints, leading to the destruction of affected joints. As most patients with RA suffer from involvement of the joints of the hands, manual ability is compromised, thereby inducing significant disability in daily living activities.[1]

Over the past decade, questionnaires and health status measures have become widely used as outcome measures in clinical trials. The Stanford Health Assessment Questionnaire (HAQ) was developed to assess patient's physical function in daily life activities.[2] Although the HAQ includes items dealing with upper limb activities, it was not originally designed to assess rheumatoid hand function. The Cochin hand functional disability scale, specifically developed for RA, is a three-factor score appropriate for descriptive purposes but not ideal for assessing changes in manual ability after medical intervention.[3] Another instrument, the Disability of the Arm, Shoulder and Hand (DASH) questionnaire, has been used in several orthopaedic and rheumatological conditions.[4] However, the DASH questionnaire includes items relating to symptoms and the patient's social involvement, and does not specifically address the patient's manual ability. Furthermore, a patient might improve on one factor but become less able on another, thereby confounding the interpretation. Therefore, to allow for quantitative comparisons, functional recovery must be assessed on a unidimensional scale.[5]

"Manual ability" may be defined as the capacity to manage daily activities requiring the use of the upper limbs, whatever the strategies involved. It can be inferred from the patient's perceived difficulty in performing activities, as determined by questionnaires.[6] A linear measure of manual ability can then be estimated from raw scores according to measurement models,[5] the most promising being the Rasch model.[7] Provided that the behavioural data fit the model requirement of unidimensionality, the manual ability measure for each patient is estimated on a linear scale defined by the difficulty of the manual activities. Once the scale is established, it is necessary to verify that the hierarchy of activity difficulties is the same for patients with different impairments.

In a preliminary study, the ABILHAND questionnaire was developed using the Rasch methodology and allowed a manual ability construct to be defined in a sample of 18 patients with RA after wrist arthrodesis.[6] This questionnaire was further validated in a larger sample of patients with stroke and found to correlate with upper limb function.[8]

The objective of the current study was to adapt and validate the ABILHAND questionnaire as a measure of manual ability in a broader sample of patients with RA and to address its test–retest reliability and sensitivity to change.

METHODS
Subjects
In total, 112 consecutive patients with RA were recruited between December 2004 and January 2005 from our rheumatology outpatient clinic. Rheumatoid arthritis was defined according to the classification criteria of the American College of Rheumatology.[9] In order to be included in the study, patients had to be on a stable medical treatment over the previous 6 months, to understand and speak French, and to present no sensorimotor deficit other than those related to RA. Disease

....

duration, number of disease-modifying anti-rheumatic drugs (DMARDs) used, tender and swollen joint counts, disease activity (28-point Disease Activity Score-C reactive protein (DAS28-CRP))[10] and radiological score[11] were used as independent indices in the validation analysis. The sample description is provided in table 1. Six additional consecutive patients with active RA requiring tumour necrosis factor (TNF) blockade were also recruited.

Patient assessment
Clinical data
Data on demographics, rheumatic treatment and handedness were collected from the patients' medical charts. All patients were evaluated in the outpatient clinic by the same investigator. The disease activity was calculated using the DAS 28-CRP.[10] The Ritchie Index was used to conduct additional analysis on the 17 joints of each upper limb.[12] An upper limb was considered to be clinically affected if at least one joint was tender or swollen. The radiological score was determined according to the simplified Sharp/van der Heijde radiological method.[11] Functional ability was assessed with the HAQ,[2] providing a score ranging from 0 (no disability) to 3 (complete disability).

Most patients were accustomed to questionnaire assessments. Nevertheless, written instructions were given to all patients on how to fill in the questionnaires prior to each test. The questionnaires were self-completed by the patients in the waiting room before the medical visit. Thorough explanations were provided to the patients as needed.

Upper limb impairments
The maximum voluntary grip and key pinch strengths were measured with the Jamar dynamometer and the pinch gauge, respectively, according to the procedure described by Mathiowetz et al.[13] Manual and digital dexterity were assessed with the Box and Block Test[14] and the Purdue Pegboard Test, respectively.[15] All tests were carried out on the dominant hand (DH) and on the non-dominant hand (NDH), starting with the hand the patient reported as less affected. The scores were z-transformed with respect to normative data[13 16] in order to account for gender, age and handedness. This procedure allows the results of both hands to be expressed on a common scale. A z-score <-1 was considered a moderate impairment, a score <-2 was considered a severe impairment.

Table 1 Sample demographics (n = 112)

Gender (n)	
Male	29
Female	83
Age, years	55 (25–82)
Handedness (n)	
Right-handed	100
Left-handed	12
Disease duration, years	10.7 (0.5–44.5)
Number of DMARDs, median (range)	2 (1–6)
Joint counts, median (range)	
Dominant UL, tender	1 (0–17)
Dominant UL, swollen	1 (0–12)
Non-dominant UL, tender	1 (0–17)
Non-dominant UL, swollen	1 (0–7)
Affected UL (n)	
None	20
One UL	9 (DH) + 4 (NDH)
Both ULs	79
Disease Activity, DAS 28-CRP	3.83 (1.27–7.78)
HAQ	1.210 (0–2.875)
Radiological Score	15 (0–86)

UL, upper limb; DH, dominant hand; NDH, non-dominant hand.
Values are mean (range) unless otherwise stated.

Manual ability
Manual ability was assessed with the ABILHAND questionnaire,[6] an inventory of 56 common manual activities of daily living that the patient is asked to evaluate on a 3-point scale (0, impossible; 1, difficult; 2, easy). The ABILHAND scoring sheets can be downloaded from www.abilhand.org. An online data analysis module featuring the scale calibration in RA allows raw scores to be directly converted into linear measures of manual ability. For each question, the patient was asked to score the feeling of difficulty independently of the limb(s) actually used to do the activity, allowing for any adaptive strategy (eg, the use of wrist splints). Activities not commonly performed in the previous 3 months were not scored and were encoded as missing.

Data analysis
Measuring manual ability through the Rasch model
The ABILHAND questionnaire was analysed with the Rasch Unidimensional Measurement Models computer program (RUMM2020; RUMM Laboratory Pty Ltd, Perth, Western Australia). For all items, the response categories (0, impossible; 1, difficult; 2, easy) were analysed according to the rating scale model.[17] The model requires, within a probabilistic framework, that the patient's response to an item depends solely on the ability of the patient and the difficulty of the response categories (computed as the sum of the item difficulty and the threshold difficulties that separate each pair of successive responses). The ability and difficulty parameters are estimated by the software, together with their standard error of measurements (ie, half of the 95% confidence interval of the true value). Based on the estimated ability of the patient and difficulty of the item, the expected response of a subject to an item can be computed by the model. The similarity between the observed and expected responses to any item is reported by the software, through a standardised residual and a χ^2 fit statistic.[18] The standardised residual is sensitive to the item discrimination. Negative values indicate that the item discriminates more than expected, and are therefore more acceptable than positive values. The χ^2 fit statistic cumulates the deviations from the model expectations; a test of significance is then applied to determine whether the χ^2 is too high to be attributed to random variations. The software also reports the person separation reliability that indicates the level of measurement precision attained.

ABILHAND item selection
Starting from the original set of 56 items, indices reported from successive analyses were used to select the items that constituted the final ABILHAND scale. The following criteria were used to select the items: (1) a response rate $>80\%$, indicating that the activity is commonly realised in the sample of patients with RA; (2) a difficulty targeted to the level of ability of the sample of patients with RA, indicating that the items provide enough information to estimate the patients' ability; (3) a discrimination similar to the other items providing the same relative weight of all response categories between items; and (4) an adequate fit to the Rasch model, indicating that the items contribute to a unidimensional manual ability construct.

Determining the scale invariance through differential item functioning
Once satisfactory metric properties were achieved, the invariance in the item difficulty hierarchy among patient subgroups (eg, men vs women) was tested.[19] In total, 11 mutually exclusive differential item functioning (DIF) subgroups were formed based on the following criteria: (1) age ($\leqslant$ median age

of 56 years vs > median), (2) gender (male vs female), (3) disease duration ($\leqslant$ median duration of 8.5 years vs > median), (4) number of DMARDs ($\leqslant$ median number of two DMARDs vs > median) (5–6) cumulative tender and swollen joints on the DH and the NDH ($\leqslant$ median cumulated count of 2 for both the DH and the NDH vs > median), (7) affected side (both limbs affected vs one or less limb affected), (8) disease activity ($\leqslant$ median DAS 28-CRP of 3.72 vs > median), (9) radiological score ($\leqslant$ median radiological score of 8 vs > median), (10) physical function ($\leqslant$ median HAQ of 1.13 vs > median) and (11) manual ability ($\leqslant$ median ABILHAND measure of 1.96 logits vs > median).

Test–retest reliability

The test–retest reliability of the ABILHAND questionnaire was studied in a subgroup of 35 patients receiving a stable medical treatment who were reassessed after an average interval of 5.4 months (range 3 to 7.6). First, the invariance of item difficulty hierarchy was assessed over time. Second, the intraclass correlation coefficient (ICC) between patient measures was determined for both assessments. Functional impairments of both upper limbs were also assessed at the re-evaluation in order to determine their relationship with manual ability.

Responsiveness

A preliminary assessment of the responsiveness of the ABILHAND scale was made by comparing patient measures at two time points in two groups of patients: (1) the 35 patients receiving a stable treatment after an average interval of 5.4 months (range 3.0 to 7.7), and (2) 6 patients following an anti-TNF treatment after an average interval of 11 weeks (range 8.0 to 18.0). In order to detect a potential trend, patient measures in each group were compared by paired t test.

Construct validation

The relationship between ABILHAND and demographic and clinical indices was tested either by t test (when comparing two groups) or by correlation coefficient (for continuous predictors). The level of statistical significance was set to p<0.05 for all analyses.

RESULTS

Refinements of the ABILHAND questionnaire for patients with RA

From the original 56-item questionnaire, items were selected through successive Rasch analyses to construct the final ABILHAND RA scale. In all, 29 items were removed: 21 were too easy for patients with RA, 4 were not commonly used by the patients (response rate <80%; eg, winding up a wrist watch), 2 were significantly less discriminating than other items (eg, buttoning up a shirt), and 2 did not contribute to the definition of a unidimensional manual ability scale (eg, counting bank notes). The resulting scale comprised 27 manual activities specifically selected for patients with RA (fig 1, table 2).

Calibration of ABILHAND in RA

Figure 1 shows the resulting linear ABILHAND scale in RA. The manual ability measures are expressed in logits, a linear unit that defines the odds of successful achievement by a patient on any particular activity. Given the linear nature of the scale, at any given level of ability, a one logit difference between two patients indicates that their odds of successful achievement of any activity are in a ratio of 2.7:1 (ie, e^1:1). The zero of the scale is conventionally set at the average item difficulty.[19]

The distribution of patient measures is presented in the top panel of fig 1. Manual ability measures ranged from −2.2 to +6.4 logits (mean (SD) 1.87 (2.05)) indicating that the most able patient had an ability approximately 5430 times higher than the least able patient, as their odds of successful achievement of any particular item are in a ratio of 5430:1 (ie, $e^{8.6}$:1). This clearly illustrates the wide range of manual ability in our patient sample.

The most probable score to each item is presented as a function of the patient's manual ability measure (fig 1, middle panel). For instance, being able to easily thread a needle requires an ability of about 3 logits, whereas being able to easily pick up a can requires an ability of about 0 logits. This indicates that, whatever the ability of the patient, threading a needle is about 20 (ie, e^3) times more difficult than picking up a can for a patient with RA. Overall, patients with RA reported moderate disability, as only 10 patients with ability >4.67 logits reported that they could perform all activities easily, whereas 77 patients with ability >0.62 logits reported that they could perform all the listed activities, although with some difficulty on the most difficult ones. None of the patients reported they could not perform any of the listed activities.

The relationship between the raw score on the ABILHAND questionnaire and the linear manual ability measures was computed as the sum of the expected score to each item, expressed as a function of the underlying ability (fig 1, bottom panel). The relationship was ogival, indicating that the finite raw score range to the ABILHAND questionnaire denotes an infinite range of manual ability. Although the relationship was quasi-linear in the central scoring range, larger differences of ability were encompassed by each score point towards both ends of the scale.

The measures of perceived difficulty for the 27 retained activities in RA are presented in table 2. The activities are listed in decreasing order of difficulty (range −2.18 to 2.65 logits) from top to bottom, with higher logit values representing more difficult activities. The table also shows the standard error of the estimates of item difficulty (average 0.20 logits), the standard residual (average 0.21) and the fit statistic computed by χ^2 test. Half of the 27 retained items commonly require the use of both hands. These items tend to define the most demanding end of the manual ability scale, though some unimanual items inducing high constraints on the joints were also perceived as more challenging. Most of the items fitted the requirement of the model indicating that they contribute to the definition of a unidimensional variable. Overall, the person separation reliability of the final scale in our sample of patients with RA was 0.95, indicating that the measurement precision is sufficiently high to separate patients in >6 strata of markedly different abilities.[20]

Differential item functioning

To test whether the difficulty hierarchy of the 27 retained activities was similar across subgroups of patients with RA, we defined dichotomous categories of patients according to 11 splitting criteria. The perceived difficulty of each item was contrasted between each mutually exclusive subgroup in the differential item functioning plots of fig 2. As most activities lay within the 95% confidence interval of the identity line, the items appeared to be ranked at the same relative difficulty across demographic and clinical subgroups of patients. There were a few exceptions. For instance, items such as (b) screwing on a nut, (e) hammering a nail and (m) handling a stapler presented greater difficulty for women than for men. Such activities require more strength and are more commonly performed by men, suggesting that the lack of invariance might reflect cultural factors.
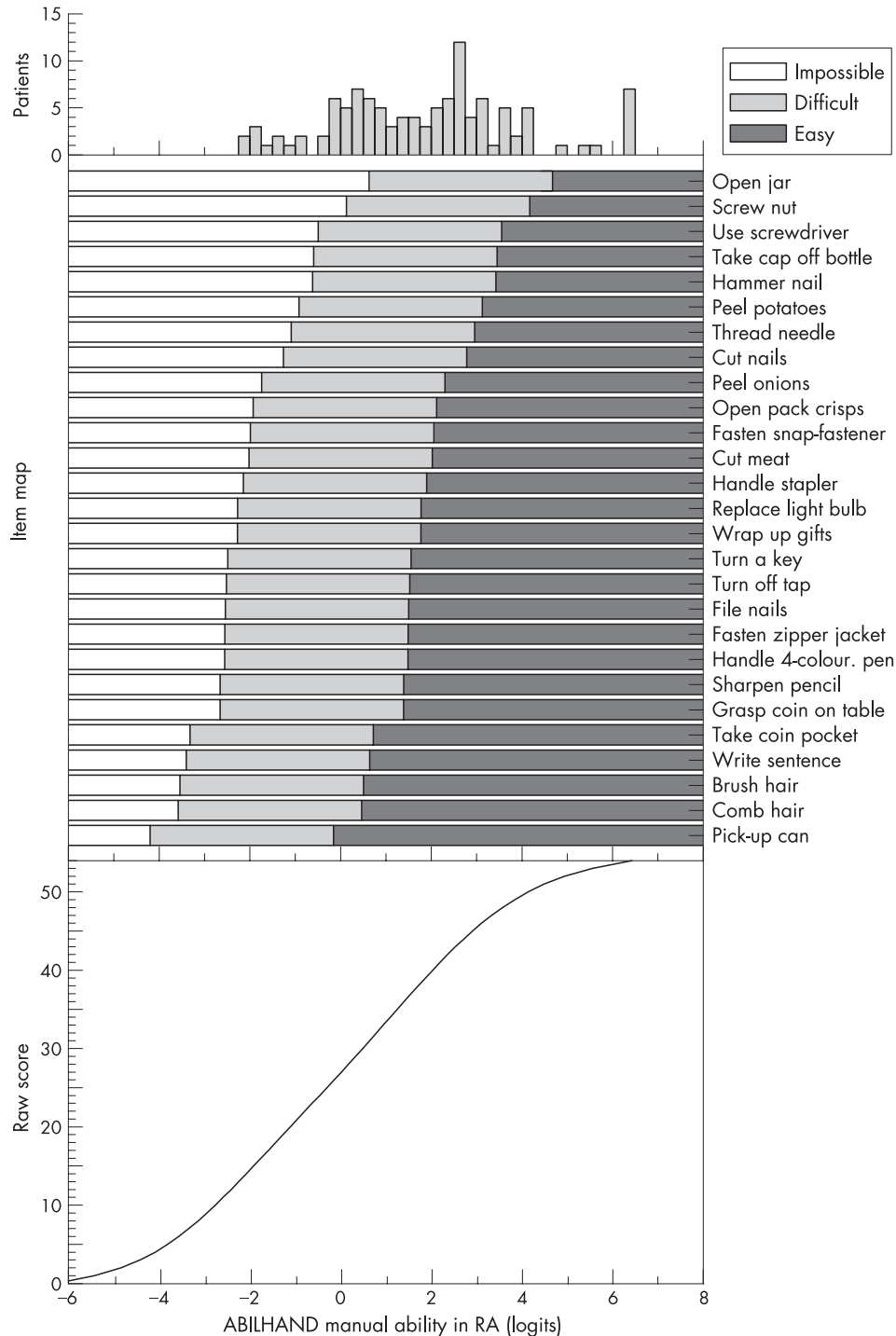
Figure 1 Structure of the ABILHAND manual ability variable in RA. The structure of the ABILHAND manual ability construct in RA is shown via the patient measures distribution (topl), the relationship between raw scores and the linear manual ability measure (bottom) and the item map (middle), which presents the evolution of the most probable score (either impossible, difficult or easy) on each item as a function of the patient's manual ability measure. The threshold measures between consecutive response categories are located at −2.06 and +2.06 logits from the difficulty of each item. The raw score is obtained by computing the sum of the expected score to each item as a function of the patient's manual ability. The non-linear relationship between raw scores and logit measures shows that larger changes of ability are denoted by the same change in raw score outside the central scoring range.

Test–retest reliability

Test–retest reliability was assessed by comparing the item difficulty hierarchy perceived by 35 patients receiving a stable medical treatment for RA. As shown in fig 2 (bottom right panel), all items lay within the confidence interval, indicating that the hierarchy of item difficulty was invariant over an average interval of 5.4 months. Patient measures at both assessments were therefore expressed on the same manual ability scale. The average ABILHAND measures (1.61 (1.83) logits at t1; 2.07 (1.81) logits at t2) did not significantly differ between assessments of patients receiving a stable treatment (paired t test, t = 1.82, p = 0.08). Moreover, the ICC between patient measures obtained at both assessments (0.74) indicates that the ABILHAND test–retest reliability is good.[21]

Table 2 ABILHAND calibration for rheumatoid arthritis patients (n = 112)

Items	Usually bimanual	Difficulty (logits)	SE (logits)	Residual (z)	Fit* (χ^2)	p Value
a. Opening a screw-topped jar	Yes	2.65	0.21	−0.40	1.57	0.67
b. Screwing on a nut		2.15	0.26	−0.46	1.90	0.59
c. Using a screwdriver		1.53	0.23	−1.83	2.73	0.43
d. Taking the cap off a bottle	Yes	1.43	0.21	−1.06	7.11	0.07
e. Hammering a nail	Yes	1.40	0.25	−1.20	0.51	0.92
f. Peeling potatoes with a knife	Yes	1.11	0.21	0.48	6.01	0.11
g. Threading a needle	Yes	0.93	0.22	−0.37	1.73	0.63
h. Cutting one's nails	Yes	0.76	0.21	−2.38	1.89	0.60
i. Peeling onions	Yes	0.28	0.22	−0.80	7.02	0.07
j. Tearing open a pack of crisps	Yes	0.10	0.21	0.72	5.75	0.12
k. Fastening a snap-fastener (eg, jacket, bag)	Yes	0.03	0.21	−0.61	3.05	0.38
l. Cutting meat	Yes	0.00	0.21	−1.96	6.51	0.09
m. Handling a stapler		−0.12	0.22	0.52	0.20	0.98
n. Replacing a light bulb		−0.25	0.24	0.87	4.18	0.24
o. Wrapping up gifts	Yes	−0.25	0.23	−0.44	2.75	0.43
p. Turning a key in a keyhole	Yes	−0.47	0.22	0.39	2.73	0.43
q. Turning off a tap		−0.50	0.22	0.30	5.24	0.16
r. Filing one's nails		−0.52	0.23	−0.20	1.10	0.78
s. Fastening the zipper of a jacket	Yes	−0.53	0.22	−2.01	1.88	0.60
t. Handling a 4-colour ballpoint pen with one hand		−0.54	0.24	2.15	4.20	0.24
u. Sharpening a pencil	Yes	−0.63	0.24	−0.72	1.17	0.76
v. Grasping a coin on a table		−0.63	0.22	1.40	5.86	0.12
w. Taking a coin out of the pocket		−1.30	0.24	0.72	6.52	0.09
x. Writing a sentence		−1.38	0.24	0.26	1.16	0.76
y. Brushing one's hair		−1.52	0.24	0.22	0.18	0.98
z. Combing one's hair		−1.56	0.25	0.61	0.16	0.98
aa. Picking up a can		−2.18	0.28	−0.40	1.50	0.68

*Degrees of freedom = 3.

Responsiveness

A preliminary assessment of the responsiveness of the ABILHAND questionnaire was made by evaluating score changes in six patients before and after treatment with TNF blockers. Patient manual ability measures, HAQ scores and disease activities at two time points were contrasted for these six patients (fig 3; filled symbols), and for 35 control patients with RA receiving a stable treatment (open symbols). Patients on stable therapy did not significantly improve their manual ability on either the DAS28-CRP or HAQ. In contrast, paired t tests showed a significant improvement in manual ability (p = 0.03) and significant decreases in DAS28-CRP (p = 0.01) and HAQ score (p = 0.02) in patients given TNF blocking agents.

Construct validation

Patients' upper limb impairments on the dominant hand (DH) and on the non-dominant hand (NDH) were evaluated at reassessment for 35 patients. Grip strength (mean (SD) −2.31 (1.44) on the DH and −1.87 (1.25) on the NDH), key pinch strength (−1.71 (1.85) on the DH and −1.41 (1.90) on the NDH) and manual dexterity (−2.30 (1.51) on the DH and −2.36 (1.18) on the NDH) were the most commonly impaired functions: two-thirds of the patients presented moderate impairment and half of them had severe impairment. Patients had less impaired digital dexterity (−1.46 (1.83) on the DH and −1.11 (1.49) on the NDH): only 50% of them presented moderate impairment and a third had severe impairment.

The relationship of demographic, clinical and functional variables with manual ability is presented in table 3. No significant relationship was found between ABILHAND measures and demographic indices (gender, age and handedness). Manual ability measures were significantly related to clinical indices (affected hands, disease duration, number of DMARDs, joint counts, disease activity and radiological score) and to functional ability as assessed by the HAQ. This indicates that the ABILHAND scale discriminates between patients with different disease severity. Manual ability was also moderately

related to grip and key pinch strength and to digital dexterity on both hands. A weaker relationship was found with manual dexterity on both hands.

DISCUSSION

Questionnaires for functional disability in RA are increasingly used to evaluate disease progression and responsiveness to therapy, such as TNF blockers.[22] [23] This study was a follow-up to a preliminary investigation of 18 patients with RA.[6] The objective of the current study was to validate the ABILHAND questionnaire as a measure of manual ability in a large cohort of 112 patients with RA and to address its test–retest reliability and sensitivity to change. From the original pool of 56 items, successive Rasch analyses showed that 27 items, that generate higher constraints in the upper limb joints were more demanding and capable of discriminating patients' manual ability. The resulting ABILHAND scale cover all the domains of the original 56 items questionnaire: communication, cooking, dressing, feeding, grooming and miscellaneous activities. The 27 retained items define a unidimensional (all p values >0.05, table 2) and linear scale, allowing recovery of manual ability to be assessed quantitatively over time within and between patients.

The resulting manual ability scale is also consistent with the type of manual involvement across different diagnoses. Thus, items generating higher compression forces in the upper limb joints are perceived as more difficult for patients with RA than for patients with stroke.[8] This indicates that the 27 selected items are specifically targeted to the functional limitations of unselected patients with RA. When compared with the items retained in chronic stroke patients, it is striking that (1) more unimanual activities have been retained in RA and (2) some activities are perceived with different difficulty in both diseases. For instance, three activities are more difficult for patients with chronic stroke because they involve greater bimanual coordination (wrapping up gifts, filing one's nails and cutting meat) and two are more difficult for patients with RA because they
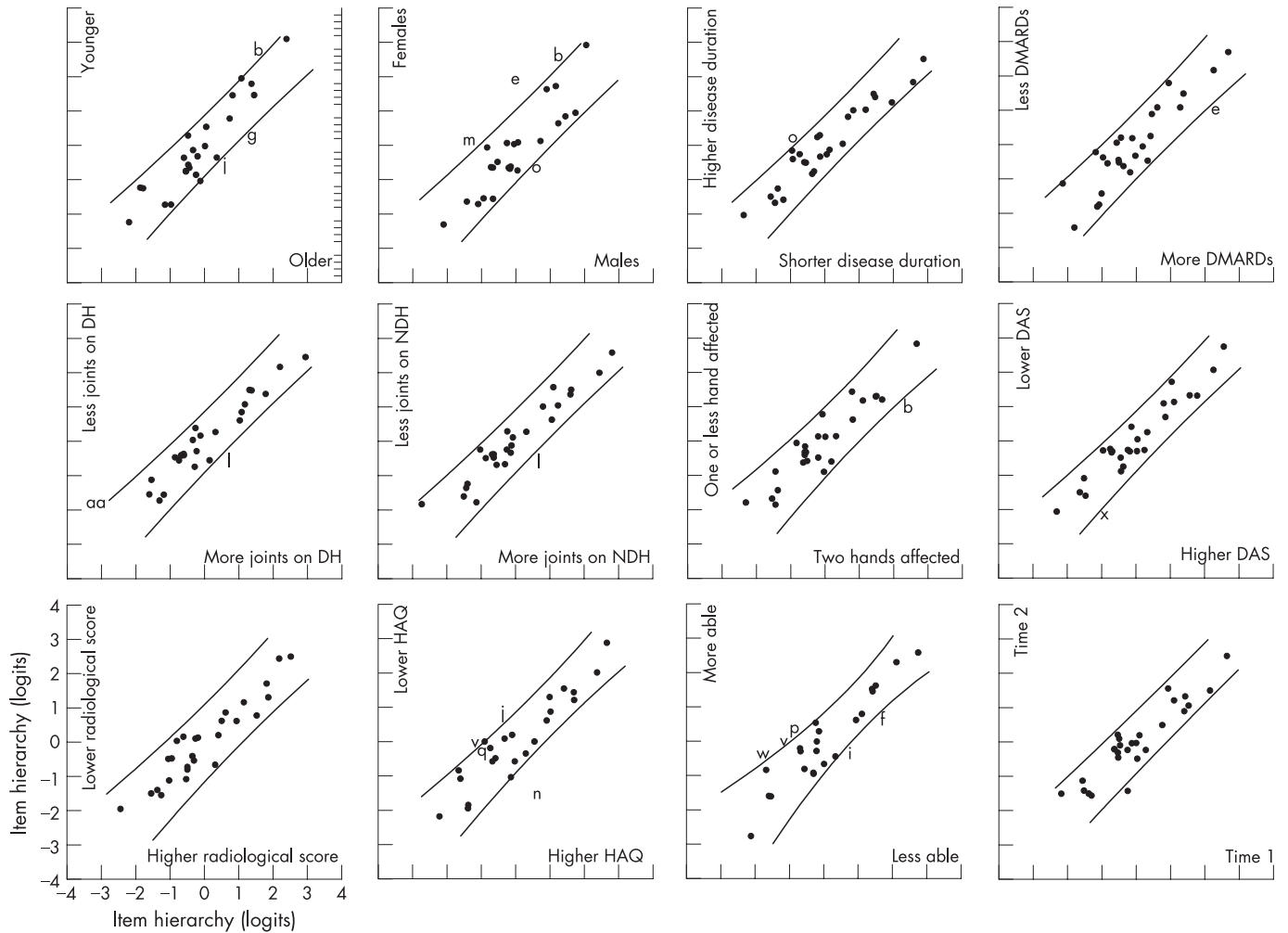
Figure 2 Invariance of the ABILHAND scale in subgroups of patients with rheumatoid arthritis. In each panel, the item difficulty hierarchy (dots) was compared in dichotomous subgroups of patients, divided according to 11 splitting criteria and over time (bottom right panel). More difficult items are shown in the top right part of each panel. Control lines (solid lines) indicate the 95% confidence intervals of the ideal invariance. Items lying within the control lines were ranked with the same hierarchy in both subgroups of patients. Outliers are identified by the labels shown in table 2. The item difficulty hierarchy was compared within the sample of 112 patients; the test–retest invariance of the scale (bottom right panel) was compared in 35 patients receiving a stable medical treatment after an average interval of 5.4 months (ICC = 0.74).

generate higher mechanical constraints in the upper limb joints (taking the cap off a bottle, fastening a snap-fastener).

The difficulty of ABILHAND items appears to be perceived invariantly between subsamples of patients with different demographic and clinical conditions. Although marginal outliers were identified for a few criteria (eg, gender, HAQ) specific data analysis protocols, including a larger sample of patients, could be used in the future to confirm this observation on an
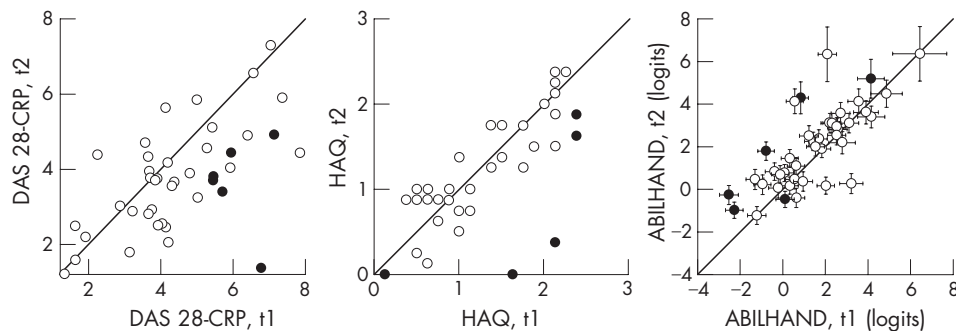

Figure 3 Reassessment of 28-point Disease Activity Score-C reactive protein (DAS-28 CRP) scale, Health Assessment Questionnaire (HAQ) and ABILHAND. Reassessment of disease activity (left panel), HAQ score (middle panel) and ABILHAND measure (right panel) after an interval of 5.4 months for patients receiving a stable medical treatment (open symbols) and an interval of 11 weeks for patients receiving tumour necrosis factor (TNF) blockers (filled symbols). Error bars in the ABILHAND plot represent 1SE of measurement. Points lying on the identity line (straight diagonal) represent no change. Patients on stable treatment showed a slight, although not significant, tendency towards improvement of DAS28-CRP and ABILHAND measure. Patients on anti-TNF treatment showed a significant improvement in DAS, HAQ and ABILHAND.

Table 3 Relationship of manual ability to demographic and clinical variables (n = 112)

Variable	Statistic*	p Value
Gender	t = −1.588	0.12
Age	R = −0.123	0.12
Handedness	t = 0.979	0.33
Affected hand(s)	t = −4.160	<0.001
Disease duration	R = −0.195	0.04
Number of DMARDs	ρ = −0.281	<0.01
Joint counts		
Tender or swollen on DH	ρ = −0.327	<0.001
Tender or swollen on NDH	ρ = −0.397	<0.001
Disease activity	ρ = −0.403	<0.001
Radiological score	ρ = −0.267	0.03
HAQ	ρ = −0.784	<0.001
Upper limb function on DH†		
Grip strength	R = 0.581	<0.001
Key pinch strength	R = 0.422	0.01
Manual dexterity	R = 0.351	0.04
Digital dexterity	R = 0.559	<0.001
Upper limb function on NDH†		
Grip strength	R = 0.542	<0.001
Key pinch strength	R = 0.464	<0.01
Manual dexterity	R = 0.250	0.15
Digital dexterity	R = 0.534	<0.001

DH, dominant hand; NDH, non-dominant hand.
*Reported statistics are: t for Student's t tests, R for Pearson correlation coefficients and ρ for Spearman correlations; †n = 35.

item-by-item basis. Nevertheless, the observed invariance supports the clinical application of the ABILHAND questionnaire in RA.

The current study confirms that patient ability is significantly related to disease activity, disease duration and radiological damage in RA.[24] Other factors such as age, gender and handedness did not significantly influence the patients' perceived manual ability. Patients with severe upper limb impairments reported a lower perceived ability as measured with the ABILHAND questionnaire and with the HAQ, as reported by previous studies.[25][26] Manual ability is mostly related to grip strength (table 3), and this result confirms that grip strength is the most related to rheumatoid hand disability.[27] Manual ability is also related to digital dexterity. Although digital dexterity was less affected than manual dexterity in our sample, the results indicate that a small impairment in this function is significantly related to a decrease in manual ability. Indeed, activities perceived as the most difficult require fine finger movements, suggesting that digital dexterity contributes to the achievement of manual activities.

Although upper limb impairments are significantly related to manual ability, patient perception of difficulty in achieving activities is correlated to disease activity but might be influenced by other factors (eg, motivation, psychosocial factors).[23] The ABILHAND questionnaire is a behavioural measure of the patient's ability to achieve manual activities whatever the strategies involved. Unlike most functional tests, it does not rely on a standardised performance but is rather a person-centred measure of patients' perceptions in their own environment. As the retained items fit the requirements of a unidimensional measure, the scale can be easily used to monitor patient's status quantitatively using activities that are meaningful to the patient.

The ABILHAND scale could be used to preliminary evaluate patient progress after treatment with TNF blockers. The clinical efficacy of TNF blockers, shown by improvement on the DAS 28-CRP, was paralleled by the improvement of the ABILHAND manual ability measures in our small cohort of patients, suggesting that the scale shows a good responsiveness. Patients treated with these agents reported a significant improvement after an average interval of 11 weeks, whereas patients on a stable treatment did not significantly progress over the same period. Although the responsiveness of the ABILHAND needs to be confirmed in a larger sample, these preliminary results support the clinical application of the scale in longitudinal RA studies. Interestingly, the HAQ showed a small floor and ceiling effect in these patients. These limitations were previously reported and have contributed to several modifications of the HAQ.[28][29] Our results indicate that, although both scales have some activities in common, the ABILHAND has a wider range of measurement for manual ability. Therefore, it is more appropriate in longitudinal studies to discriminate small changes at both extremities of the scale. However, 61% of the variance in patient response was common to both scales, indicating that they address complementary constructs and could be used in parallel in further clinical studies.

The ABILHAND scale presents a good person separation reliability in our sample (R = 0.95) indicating that the item discrimination allows patients to be stratified into >6 groups of significantly differing abilities along the resulting manual ability scale. The 27 items are well targeted to the ability of the patients, yet cover a wide range of functional status. The observed invariance of the item hierarchy across demographic and clinical subgroups of patients supports the clinical application of the ABILHAND scale as a measure of manual ability in a wide spectrum of patients with RA. The observed invariance over time also supports the test–retest reliability of the ABILHAND manual ability measures (fig 2). Some limitations of this study could be addressed in future research, such as the qualitative investigation of the items' content in order to evaluate the clinical relevance and the marginal differential functioning of the items, and the analysis of a larger and broader RA population to verify the responsiveness of the scale.

In conclusion, the ABILHAND scale is valid, invariant across patients' subgroups and appears to be sensitive to medical treatment of RA. ABILHAND can be used to assess and follow patient recovery of manual ability, as its item difficulty hierarchy delineates the expected pattern of recovery in patients with RA. Moreover, the questionnaire is extremely easy to administer as it can be self-completed in 5 minutes by the patient in the waiting room. Future applications of the scale include the assessment of medical and/or rehabilitative procedures specifically focused on the upper limb(s) of patients with RA.

....

Authors' affiliations
Patrick Durez, Frédéric Houssiau, Department of Rheumatology, Cliniques Universitaires St-Luc, Université Catholique de Louvain, Brussels, Belgium
Virginie Fraselle, Jean-Louis Thonnard, Henri Nielens, Massimo Penta, Department of Physical Medicine and Rehabilitation, Cliniques Universitaires St-Luc, Université Catholique de Louvain, Brussels, Belgium

REFERENCES
1 **Dellhag B**, Bjelle A. A five-year followup of hand function and activities of daily living in rheumatoid arthritis patients. *Arthritis Care Res* 1999;**12**:33–41.
2 **Pincus T**, Summey JA, Soraci SA Jr, Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford health assessment questionnaire. *Arthritis Rheum* 1983;**26**:1346–53.
3 **Duruöz MT**, Poireaudeau S, Fermanian J, Menkes C, Amor B, Dougados M, Revel M. Development and validation of a rheumatoid hand functional disability scale that assesses functional handicap. *J Rheumatol* 1996;**23**:1167–72.
4 **Navsarikar A**, Gladman DD, Husted JA, Cook RJ. Validity assessment of the disabilities of the arm, shoulder, and hand questionnaire (DASH) for patients with psoriatic arthritis. *J Rheumatology* 1999;**26**:2191–4.

5 **Wright BD**, Linacre JM. Observations are always ordinal; measurement, however, must be interval. *Arch Phys Med Rehabil* 1989;**70**:857–60.
6 **Penta M**, Thonnard J-L, Tesio L. ABILHAND: a Rasch-built measure of manual ability. *Arch Phys Med Rehabil* 1998;**79**:1038–42.
7 **Rasch G**. *Probabilistic models for some intelligence and attainment tests.* Chicago: MESA Press, 1980, Originally published in 1960 by the Danish Institute of Educational Research.
8 **Penta M**, Tesio L, Arnould C, Zancan A, Thonnard JL. The ABILHAND questionnaire as a measure of manual ability in chronic stroke patients: Rasch-based validation and relationship to upper limb impairment. *Stroke* 2001;**32**:1627–34.
9 **Arnett FC**, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;**31**:315–24.
10 **Prevoo ML**, van't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;**38**:44–8.
11 **van der Heijde D**, Dankert T, Nieman F, Rau R, Boers M. Reliability and sensitivity to change of a simplification of the Sharp/van der Heijde radiological assessment in rheumatoid arthritis. *Rheumatology* 1999;**38**:941–7.
12 **Ritchie DM**, Boyle JA, McInnes JM, Jasani MK, Dalakos TG, Grieveson P, et al. Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. *Q J Med* 1968;**37**:393–406.
13 **Mathiowetz V**, Weber K, Volland G, Kashman N. Reliability and validity of pinch and grip strength evaluations. *J Hand Surg* 1984;**9A**:222–6.
14 **Mathiowetz V**, Volland G, Kashman N, Weber K. Adult norms for the Box and Block test of manual dexterity. *Am J Occup Ther* 1985;**39**:386–91.
15 **Tiffin J**, Asher EJ. The Purdue Pegboard: norms and studies of reliability and validity. *J Appl Psychol* 1948;**32**:234–247.
16 **Mathiowetz V**, Kashman N, Volland G, Weber K, Dowe M, Rogers S. Grip and pinch strength: normative data for adults. *Arch Phys Med Rehabil* 1985b;**66**:69–74.
17 **Andrich D**. A rating formulation for ordered response categories. *Psychometrika* 1978;**43**:561–73.
18 **Andrich D**, Sheridan BE. *Interpreting RUMM2020: Part II. Polytomous data,* 3rd ed. Perth, WA: RUMM laboratory Pty Ltd, 2005.
19 **Wright BD**, Stone MH. *Best test design.* Chicago: Mesa Press, 1979.
20 **Wright BD**, Masters GN. *Rating SCALE ANALYsis.* Chicago: Mesa Press, 1982.
21 **Fleiss JL**. *Statistical methods for rates and proportions.* New York: Wiley, 1981.
22 **Keystone EC**, Kavanaugh AF, Sharp JT, Tannenbaum H, Hua Y, Teoh LS, et al. Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled, 52-week trial. *Arthritis Rheum* 2004;**50**:1400–11.
23 **Klareskog L**, van der Heijde D, de Jager JP, Gough A, Kalden J, Malaise M, et al. Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial. *Lancet* 2004;**363**:675–81.
24 **Drossaers-Bakker KW**, Zwinderman AH, Vlieland TP, Van Zeben D, Vos K, Breedveld FC, et al. Long-term course and outcome of functional capacity in rheumatoid arthritis: the effect of disease activity and radiologic damage over time. *Arthritis Rheum* 1999;**42**:1854–60.
25 **Dellhag B**, Burckhardt CS. Predictors of hand function in patients with rheumatoid arthritis. *Arthritis Care Res* 1995;**8**:16–20.
26 **Vliet Vlieland TP**, van der Wijk TP, Jolie IM, Zwinderman AH, Hazes JM. Determinants of hand function in patients with rheumatoid arthritis. *J Rheumatol* 1996;**23**:835–40.
27 **Spiegel JS**, Paulus HE, Ward NB, Spiegel TM, Leake B, Kane RL. What are we measuring? An examination of walk time and grip strength. *J Rheumatol* 1987;**14**:80–6.
28 **Tennant A**, Hillman M, Fear J, Pickering A, Chamberlain MA. Are we making the most of the Stanford Health Assessment Questionnaire? *Br J Rheumatol* 1996;**35**:574–8.
29 **Wolfe F**, Michaud K, Pincus T. Development and validation of the Health Assessment Questionnaire II. *Arthritis Rheum* 2004;**50**:3296–3305.