

# Instrumented BASFI (iBASFI) Shows Promising Reliability and Validity in the Assessment of Activity Limitations in Axial Spondyloarthritis

Thijs Willem Swinnen, Milica Milosevic, Sabine Van Huffel, Wim Dankaerts, Rene Westhovens, and Kurt de Vlam

**ABSTRACT. Objective.** The Bath Ankylosing Spondylitis Functional Index (BASFI) is the most popular method to assess activity capacity in axial spondyloarthritis (axSpA), to our knowledge. It is endorsed by the Assessment of Spondyloarthritis international Society. But it may have recall bias or aberrant self-judgments in individual patients. Therefore, we aimed to (1) develop the instrumented BASFI (iBASFI) by adding a body-worn accelerometer with automated algorithms to performance-based measurements (PBM), (2) study the iBASFI's core psychometric properties, and (3) reduce the number of iBASFI items.

**Methods.** Twenty-eight patients with axSpA wore a 2-axial accelerometer while completing 12 PBM derived from the BASFI. A chronometer and both manual and "automated algorithm-based" acceleration segmentation identified movement time. Test-retest trials and methods (algorithm vs manual segmentation/chronometer/BASFI) were compared with ICC, standard error of measurement [percentage of movement time (SEM%)], and Spearman  $\rho$  correlation coefficients. Linear regression identified the optimal set of reliable iBASFI PBM.

**Results.** Good to excellent test-retest reliability was found for 8/12 iBASFI items (ICC range 0.812–0.997, SEM range 0.4–30.4%), typically with repeated and fast movements. Automated algorithms excellently mimicked manual segmentation (ICC range 0.900–0.998) and the chronometer (ICC range 0.878–0.998) for 10/12 iBASFI items. Construct validity compared with the BASFI was confirmed for 7/12 iBASFI items ( $\delta$  range 0.504–0.755). Together, sit-to-stand speed test (stBeta 0.483), cervical rotation (stBeta –0.392), and height (stBeta –0.375) explained 59% of the variance in the BASFI ( $p < 0.01$ ).

**Conclusion.** The proof-of-concept iBASFI showed promising reliability and validity in measuring activity capacity. The number of the iBASFI's PBM may be minimized, but further validation in larger axSpA cohorts is needed before its clinical use. (First Release June 15 2016; J Rheumatol 2016;43:1532–40; doi:10.3899/jrheum.150439)

## Key Indexing Terms:

ANKYLOSING SPONDYLITIS	PHYSICAL ACTIVITY	MOBILITY LIMITATIONS
QUESTIONNAIRE	VALIDITY AND RELIABILITY	TECHNOLOGY ASSESSMENT

Typical features of axial spondyloarthritis (axSpA) include inflammation of the spinal and sacroiliac joints, often accompanied with syndesmophytes, bony enthesal spurs, and joint ankylosis<sup>1</sup>. Axial inflammation and bone formation manifest clinically as inflammatory back pain, loss of spinal mobility,

and spinal stiffness<sup>2</sup>. Peripheral arthritis and enthesitis, as well as extraskelatal features such as psoriasis, uveitis, and gut inflammation may add to the systemic characteristic of the disease<sup>3</sup>.

In 1999, the Assessment of SpondyloArthritis interna-

*From the Division of Rheumatology, University Hospitals Leuven; Skeletal Biology and Engineering Research Center, Department of Development and Regeneration, KU Leuven; Musculoskeletal Rehabilitation Research Unit, Department of Rehabilitation Sciences, KU Leuven; Stadius Centre for Dynamical Systems, Signal Processing and Data Analytics, Department of Electrical Engineering, KU Leuven; iMinds, Medical Information Technology, KU Leuven, Leuven, Belgium.*

*T.W. Swinnen, PT, MSc, Doctoral Research Fellow, Division of Rheumatology, University Hospitals Leuven, and Skeletal Biology and Engineering Research Center, Department of Development and Regeneration, KU Leuven, and Musculoskeletal Rehabilitation Research Unit, Department of Rehabilitation Sciences, KU Leuven; M. Milosevic, MSc Eng, Doctoral Research Fellow, Stadius Centre for Dynamical Systems, Signal Processing and Data Analytics, Department of Electrical Engineering, KU Leuven, and iMinds, Medical Information Technology, KU Leuven; S. Van Huffel, MSc Eng, PhD, Full Professor Biomedical*

*Data Processing, Stadius Centre for Dynamical Systems, Signal Processing and Data Analytics, Department of Electrical Engineering, KU Leuven, and iMinds, Medical Information Technology, KU Leuven; W. Dankaerts, PT, PhD, Professor Rehabilitation Sciences, Musculoskeletal Rehabilitation Research Unit, Department of Rehabilitation Sciences, KU Leuven; R. Westhovens, MD, PhD, Full Professor Rheumatology, Division of Rheumatology, University Hospitals Leuven, and Skeletal Biology and Engineering Research Center, Department of Development and Regeneration, KU Leuven; K. de Vlam, MD, PhD, Principal Investigator Clinical Rheumatology, Division of Rheumatology, University Hospitals Leuven, and Skeletal Biology and Engineering Research Center, Department of Development and Regeneration, KU Leuven.*

*Address correspondence to Dr. K. de Vlam, Herestraat 49, Box 7003/13, 3000 Leuven, Belgium. E-mail: kurt.devlam@uzleuven.be*

*Accepted for publication April 19, 2016.*

tional Society (ASAS) expert group developed a “minimal core set” of 9 health-related domains to monitor all aspects of disease outcome in patients with axSpA in trial and clinical record-keeping settings<sup>4</sup>. Strong consensus was reached for the inclusion of the physical function domain across settings and the contemporary Bath Ankylosing Spondylitis Functional Index (BASFI) was recommended<sup>5,6</sup>. Over the following decade, the importance of physical function was further evidenced by its pivotal involvement in the ASAS20, ASAS40, and ASAS5/6 improvement criteria<sup>7,8</sup>.

Lacking any operational definition, the content of the physical function domain in axSpA remained surprisingly ill-defined until ASAS attempted in 2010 to also integrate the effect of disease into the outcome assessment by tailoring the World Health Organization’s (WHO) International Classification of Functioning, Disability and Health (ICF) to develop “core sets” for the evaluation of axSpA<sup>9</sup>. Physical function is now largely reflected in the ICF components “activities” and “participation,” where “activities” are defined as “the execution of tasks or actions by an individual” and “participation” as “involvement in life situations”<sup>10,11</sup>.

Applying the ICF clinically, ASAS/WHO has proposed a patient-reported “capacity and performance qualifier” (ranging from 0 = no difficulty to 4 = complete difficulty to execute a task) to evaluate “activity limitations and participation restrictions”<sup>12</sup>. Similarly, the recently developed ASAS Health Index adopted patient-reported dichotomous (I agree/I do not agree) response options to evaluate the broader construct of “functioning”<sup>13</sup>. Unfortunately, this first implementation of the WHO/ASAS/ICF axSpA core sets<sup>9,12</sup> has failed to recognize the crucial distinction between what a person can do in a standardized environment (activities: capacity qualifier) versus in a real-life situation (participation: performance qualifier)<sup>10</sup>. This is problematic for valid outcome assessment in clinical care and research. Indeed, to improve an individual’s intrinsic activity capacity, a rehabilitation approach focusing more on body functions and structures is needed, while contextual rehabilitation goals targeting environmental (e.g., ergonomic adaptations) and personal (e.g., motivation) factors may add to performance in the real-life environment. Similar to the popular BASFI, the WHO/ASAS/ICF axSpA core sets still rely on the patient’s self-report and target the difficulty experienced with the execution of activities in daily life. A large body of evidence suggests that the cognitive process of mapping this experience into the construct of activity capacity or performance is patient-specific and can be distorted by psychological factors such as pain<sup>14</sup>, differences in reference frame<sup>15</sup>, cognitive impairments, motivation, anxiety, and depression<sup>16,17</sup>.

Performance-based measures (PBM) deliver a more direct and standardized observation of activity capacity and are less influenced by psychological or environmental factors than patient-reported outcomes<sup>14,18</sup>. Typically, a trained observer translates the visual inspection of the activity of interest into

an activity capacity metric such as time (using a chronometer), number of repetitions, or distance<sup>19</sup>. The 1-week test-retest reliability of a series of common PBM reflecting each BASFI item was established in patients with axSpA, minimizing the involvement of fluctuating activity capacity<sup>20</sup>. Interestingly, PBM seem to also identify small improvements in activity capacity in patients with axSpA not responding to antitumor necrosis factor (anti-TNF) therapy<sup>21</sup>. Similar in osteoarthritis (OA), PBM show different recovery pathways after joint replacement and do not have ceiling effects in longterm followup of activity capacity<sup>22</sup>. However, especially for large clinical trials, feasibility is of major concern because trained observers, facilities, and presence of the patient are needed<sup>19</sup>.

Automated identification of movement data using sensors may have the potential to speed up data collection and downsize observer training of PBM in routine practice. Semiautomated accelerometer algorithms with proven psychometric properties exist in patients with parkinsonism and stroke, but are limited to selected tests such as walking, standing, and sit-to-stand (STS) activities<sup>23,24,25</sup>. In patients with axSpA, the use of body-worn sensors is still in its infancy and focuses on accelerometer algorithms for physical activity assessment<sup>26</sup>. Therefore, our proof-of-concept study aimed to (1) present the development of the instrumented BASFI (iBASFI), adding fully automated algorithms during PBM to obtain activity capacity outcomes in axSpA, (2) evaluate the test-retest reliability of the automated algorithms and their criterion validity in comparison with manual feature selection in axSpA, (3) compare these novel activity capacity outcomes for construct validity to chronometer-based PBM and the BASFI; and (4) create a preliminary optimal set of PBM that reflected the BASFI best.

## MATERIALS AND METHODS

**Participants.** Our psychometric study randomly included 28 subjects with a diagnosis of axSpA according to the ASAS classification criteria<sup>2</sup>, verified by an ASAS expert rheumatologist (KDV), from the outpatient axSpA clinic at the University Hospitals of Leuven, Belgium. Exclusion criteria were (1) a history of spinal fractures or other fractures within 12 months, lower quadrant musculoskeletal injuries not related to SpA, discitis, and spondylolisthesis, (2) current symptoms of severe health conditions (e.g., heart failure), and (3) not being able to stand or walk without an aid. All subjects provided written informed consent prior to participation according to the Declaration of Helsinki. The study protocol fulfilling the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) criteria<sup>27</sup> (completed COSMIN form available from the authors on request) was approved by the Medical Ethics Committee of the University Hospitals Leuven (ML 5236).

**Anthropometrics and demographics.** Anthropometric measures were taken by the same observer prior to other assessments. Height was measured with a stadiometer (Holtain Ltd.) to the nearest 0.1 cm and weight was measured with a digital scale (SECA) to the nearest 0.1 kg. Work status was assessed with the work productivity survey<sup>28</sup>.

**Patient-reported activity capacity.** Patient-reported activity capacity was assessed with the ASAS-endorsed and widely accepted BASFI questionnaire<sup>29</sup>. This instrument asks the respondent to rate the perceived difficulty

in executing commonly limited activities in axSpA on an 11-point numerical rating scale with a range of 0 (easy) to 10 (impossible)<sup>30</sup>. A total score is calculated by dividing the sum of all items by 10, producing a score between 0 and 10. The psychometric properties of the BASFI in axSpA are well established<sup>31,32,33,34</sup>.

**Performance-based activity capacity.** Performance-based activity capacity was evaluated by performance-based tests (Figure 1) mimicking the BASFI<sup>20</sup> and the WHO/ICF/ASAS core sets<sup>9</sup> activities to ensure content validity. Each performance was timed with a chronometer (Geonaute) in seconds (0.01 s). The observer received a written manual and was trained to standardize the onset and end of movement. Concurrently, a body-worn, 2-axial accelerometer (Sensewear Pro 3 Armband, Bodymedia Inc.) was mounted on the triceps region (middle between humeral head and olecranon) of the dominant arm to automatically identify acceleration vectors ( $m/s^2$ ) at 32 Hz across the longitudinal and transversal axis. For the maximal reach test, movement duration was complemented with distance using a stadiometer (Janssen-Fritzen), while for the looking-over-the-shoulder test, only range of motion collected with a mounted goniometer (ORTEC) was considered relevant. All PBM were repeated twice with 30 s in between (test-retest reliability) and most tests consisted of both a “self-selected pace” and “maximal speed” trial. The PBM were randomized and counterbalanced for side. Supplementary Data 1 (available online at jrheum.org) has a detailed description of each PBM test reflecting the candidate iBASFI items (iBASFI development set; Supplementary Data 2, available online at jrheum.org).

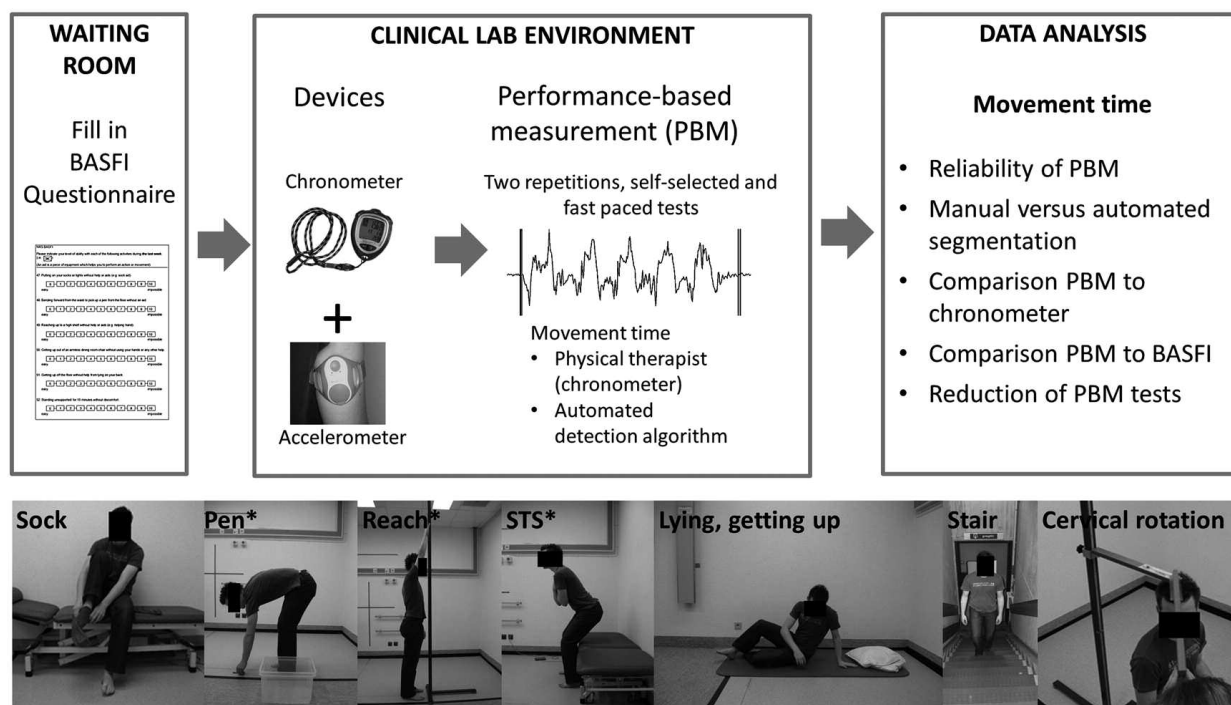
**Data reduction and statistical analysis.** Descriptive data were presented as mean, median, and 25th and 75th percentile. Normal distribution of all variables was evaluated with the Shapiro-Wilk test ( $p < 0.05$ ). Movement time was extracted from low-pass filtered accelerometer signals using custom-written automated algorithms (accelerometer<sub>auto</sub>) in MatLab (Mathworks) and by calculating the mean of manual signal segmentations (accelerometer<sub>manual</sub>) by 2 blinded evaluators (bachelor level physical therapists not involved in any part of our study). The test-independent automated

algorithm entailed a set of heuristic rules obtained from a pilot set of acceleration data ( $n = 3$ ), applied wavelet-based filtering and took into account SD of signal variables in a sliding window for both the transverse and longitudinal axis to detect the start and end of each movement.

Trials (within-session attempt 1 vs 2) and methods (accelerometer<sub>man</sub> vs accelerometer<sub>auto</sub>, chronometer vs accelerometer<sub>auto</sub>, accelerometer<sub>auto</sub> vs corresponding BASFI-item, accelerometer<sub>auto</sub> vs corresponding BASFI total) were compared with the ICC [test-retest trials: 2-way mixed; intermethod (see comparisons above): 2-way random model], the 95% CI for the standard error of measurement [SEM; also clinically expressed in percent of total movement time (SEM%)], and the Spearman  $\rho$  correlation coefficient. Criteria to evaluate ICC were  $< 0.70$  (poor),  $0.70-0.79$  (adequate),  $0.80-0.89$  (good), and  $> 0.90$  (excellent)<sup>20</sup>. We hypothesized a significant correlation of 0.50 or more to confirm convergent construct validity between PBM and the BASFI questionnaire and its items<sup>35</sup>. A stepwise linear regression was used to model the BASFI from an optimal set of reliable ( $ICC > 0.80$ ) performance-based tests obtaining a pilot version of the short iBASFI. Sex, height, weight, and body mass index were also included into the regression's independent variables to exclude anthropometric effects. The criteria to include or exclude variables were a probability of  $F \leq 0.05$  (in) and  $\geq 0.10$  (out), respectively. To avoid colinearity, the variance inflation factor was set at  $< 3$ , and the tolerance level at  $> 0.10$ .

## RESULTS

Demographics and disease-related characteristics are presented in Table 1 and indicate a typical outpatient sample of patients with axSpA in terms of age, disease-related variables, and medication use<sup>36</sup>. However, the severity of thoracic kyphosis (tragus-to-wall distance) and hip involvement (intermalleolar distance) was rather limited with median values of 1 and 2 out of 10, respectively. Of the 29



**Figure 1.** Setup for the instrumented BASFI psychometric proof-of-concept study in patients with axial spondyloarthritis ( $n = 28$ ). Top row illustrates study flow. Bottom row illustrates all PBM. \* All tests were repeated twice and included both a self-selected and a fast-paced variant. All other performance-based tests were fast-paced, except for cervical rotation. BASFI: Bath Ankylosing Spondylitis Functional Index; PBM: performance-based measurement.

Table 1. Demographic, anthropometric, and disease-specific measures for all patients with axial spondyloarthritis (n = 28). Height and weight were measured with a stadiometer (Holtain Ltd.) and digital scale (SECA), respectively.

Variables	Mean	SD	P25	Median	P75	Min	Max
Age, yrs	43.69	10.45	35.58	42.95	53.08	23.90	60.40
Disease duration, yrs*	12.45	10.52	3.60	10.15	20.23	0.00	38.70
Height, cm	169.41	11.53	160.05	170.50	177.08	148.90	193.10
Weight, kg	75.45	18.52	61.15	76.00	86.95	46.70	120.00
BMI, kg/m <sup>2</sup>	26.19	5.71	21.65	24.57	29.73	17.48	39.36
BASDAI, 0–10	3.29	1.96	1.40	2.90	4.70	0.60	7.40
BASFI, 0–10	3.14	2.19	1.65	3.15	5.15	0.00	8.10
BASMI, 0–10	3.11	1.60	1.85	3.00	3.95	0.60	7.60
Cervical rotation, 0–10	3.75	2.22	2.00	3.50	5.00	0.00	10.00
Tragus-to-wall, 0–10	1.79	1.71	1.00	1.00	2.75	0.00	9.00
Lateral flexion, 0–10	4.50	1.97	3.00	4.50	6.00	0.00	8.00
Intermalleolar, 0–10	2.46	1.84	1.00	2.00	3.75	0.00	7.00
Modified Schober, 0–10	5.68	1.94	4.25	5.00	7.00	1.00	10.00
Frequencies (%)							
Male/female	16/12 (57/43)						
NSAID, yes/no	17/11 (61/39)						
Biologicals, yes/no	15/13 (54/46)						
DMARD, yes/no	9/19 (32/68)						

\* Starting from diagnosis. P: percentile; BMI: body mass index; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; BASMI: Bath Ankylosing Spondylitis Disease Metrology Index; NSAID: nonsteroidal antiinflammatory drugs; DMARD: disease-modifying antirheumatic drugs.

patients invited, only 1 male subject refused to participate in our study because of a lack of time (0.03%).

*Test-retest reliability of the instrumented performance-based (iBASFI) tests.* Full data are shown in Table 2. Good to excellent test-retest reliability for instrumented PBM was found with the lowest ICC values for sock, reach, shoulder speed, and STS tests (ICC range 0.528–0.785), but very high ICC values for pen, pen speed, reach height, STS speed, lying down, getting up, stair climbing, and cervical rotation tests (ICC range 0.812–0.997). Overall, single and/or self-selected

pace movements were less reproducible than repeated and/or fast-paced movements (e.g., pen ICC 0.812 vs pen speed ICC 0.974). SEM% values ranged from 0.4% to 23.9%, excluding the pen (30.4%) and reach tests (32.0%).

*Automated algorithm versus manual segmentation and versus a chronometer.* Automated segmentation of acceleration signals to obtain movement duration excellently mimicked the mean value of manual segmentation (ICC range 0.900–0.998; Table 3), except for the reach (ICC 0.727) and STS tests (ICC 0.599). SEM% values ranged from 3.8% to

Table 2. Descriptives and test-retest reliability for the instrumented performance-based tests (iBASFI) in patients with axial spondyloarthritis (n = 28).

Tests	Descriptives		ICC <sup>2</sup>	Trial 1 vs Trial 2		
	Mean	SD		95% CI	SEM	% Task
Sock, s	20.180	5.880	0.765	0.482–0.893	2.850	14.125
Pen, s	6.950	4.880	0.812	0.637–0.908	2.116	30.445
Pen speed, s	16.600	9.000	0.974	0.887–0.991	1.451	8.742
Reach, s	3.800	1.770	0.528	0.185–0.753	1.216	32.001
Reach height, cm	214.696	16.689	0.997	0.993–0.999	0.914	0.426
Shoulder speed, s	8.030	1.980	0.785	0.441–0.911	0.918	11.433
STS, s	2.750	1.320	0.753	0.537–0.877	0.656	23.856
STS speed, s	16.260	5.710	0.959	0.803–0.986	1.156	7.111
Lying down, s	9.210	11.770	0.979	0.954–0.991	1.706	18.519
Getting up, s	11.090	16.630	0.993	0.985–0.997	1.391	12.546
Stair climbing, s	10.870	3.190	0.863	0.714–0.935	1.181	10.862
Cervical rotation, °	59.857	20.990	0.994	0.985–0.997	1.626	2.716

iBASFI: instrumented Bath Ankylosing Spondylitis Functional Index; SEM: standard error of measurement; % task: SEM expressed as a portion of the task's mean movement time; STS: sit-to-stand.



Table 3. Concurrent criterion (compared with manual segmentation) and convergent construct (compared with chronometer) validity of the automated algorithm in patients with axial spondyloarthritis (n = 28).

Tests	Automated Algorithm vs Manual Segmentation				Automated Algorithm vs Chronometer			
	ICC	95% CI	SEM	% Task	ICC	95% CI	SEM	% Task
Sock, s	0.917	0.829–0.960	1.694	8.395	0.878	0.217–0.965	2.054	10.177
Pen, s	0.900	0.769–0.955	1.543	22.204	0.894	0.785–0.949	1.589	22.861
Pen speed, s	0.995	0.981–0.998	0.636	3.834	0.993	0.978–0.997	0.753	4.536
Reach, s	0.727	0.063–0.906	0.925	24.337	0.532	0.133–0.768	1.211	31.865
Shoulder speed, s	0.967	0.932–0.985	0.360	4.479	0.888	0.402–0.964	0.663	8.252
STS, s	0.599	0.189–0.811	0.836	30.396	0.770	0.565–0.886	0.633	23.020
STS speed, s	0.977	0.930–0.991	0.866	5.326	0.951	0.790–0.983	1.264	7.773
Lying down, s	0.994	0.977–0.998	0.912	9.899	0.998	0.996–0.999	0.526	5.715
Getting up, s	0.998	0.996–0.999	0.744	6.706	0.997	0.993–0.999	0.911	8.213
Stair climbing, s	0.975	0.947–0.988	0.504	4.640	0.892	0.428–0.965	1.048	9.644

SEM: standard error of measurement; % task: SEM expressed as a portion of the task's mean movement time; STS: sit-to-stand.

30.4%. Thus, criterion validity of the automated algorithm in comparison to manual segmentation was confirmed for almost all tests.

In comparison with the chronometer, the automated algorithm validly assessed movement duration (ICC range 0.878–0.998) apart from the maximal reach (ICC 0.532) and STS tests (ICC 0.770). SEM% values ranged from 4.5% to 23.0%, ignoring the reach test (31.9%). Together, convergent construct validity of the automated algorithm was established for almost all tests.

**Construct validity and item reduction of the iBASFI.** Good to excellent convergent construct validity for the instrumented pen, pen speed, STS speed, lying down, getting up, stair climbing, and cervical rotation tests was found evidenced by significant and good associations with the BASFI scale (Spearman  $\rho$  range 0.504–0.638; Table 4). All these tests, except for lying down, showed similar correlations with their

corresponding BASFI item. Surprisingly, none of the timed arm tests reached significance, and none showed relevant correlations with the BASFI or its item on reaching toward a shelf, while the maximal distance reached almost showed a significant and moderate correlation with the BASFI, but not its corresponding item.

Based on the test-retest reliability and construct validity in comparison with chronometer results for the automated algorithm, the sock, pen speed, shoulder speed, STS speed, lying down, getting up, stair, and cervical rotation tests were entered in the model in addition to anthropometric variables. The core set of the iBASFI items consisted of the STS speed test, the cervical rotation test, and height (Figure 2). This core set explained 59% of the variance in the BASFI, confirming convergent construct validity of the composite and successful reduction of items. No issues on colinearity were detected.

Table 4. Convergent construct validity between instrumented and self-reported (BASFI) activity capacity in patients with axial spondyloarthritis (n = 28).

Tests	r BASFI	Hypothesis Testing	p	r Item	Hypothesis Testing	p
Sock	0.191	N	0.330	0.306	N	0.113
Pen	0.504	C	<b>0.006</b>	0.521	C	<b>0.004</b>
Pen speed	0.548	C	<b>0.003</b>	0.385	N	<b>0.043</b>
Reach	0.108	N	0.584	0.230	C	0.238
Reach height	–0.363	N	0.058	–0.203	N	0.300
Shoulder speed	0.225	N	0.250	0.265	N	0.172
STS	0.354	N	0.065	0.279	N	0.151
STS speed	0.540	C	<b>0.003</b>	0.547	C	<b>0.003</b>
Lying down	0.532	C	<b>0.004</b>	0.365	N	0.056
Getting up	0.638	C	<b>&lt; 0.001</b>	0.580	C	<b>0.001</b>
Stair climbing	0.592	C	<b>0.001</b>	0.575	C	<b>0.001</b>
Cervical rotation	–0.539	C	<b>0.003</b>	–0.755	C	<b>&lt; 0.001</b>

Significant data are in bold face. BASFI: Bath Ankylosing Spondylitis Functional Index; iBASFI: instrumented BASFI; item: BASFI item corresponding to the iBASFI performance-based test; STS: sit-to-stand; C and N: hypothesis on construct validity  $r \geq 0.50$  was confirmed (C) or not confirmed (N).

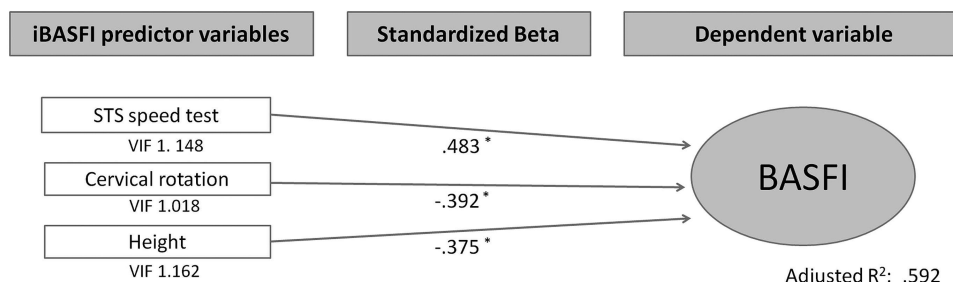


Figure 2. Stepwise modeling of self-reported physical function (BASFI) with all reliable iBASFI performance-based tests as independent variables in patients with axial spondyloarthritis (n = 28). \* p < 0.01. BASFI: Bath Ankylosing Spondylitis Functional Index; iBASFI: instrumented BASFI; STS: sit-to-stand, linear regression with stepwise selection; VIF: variance inflation factor.

## DISCUSSION

To our knowledge, our proof-of-concept study was the first to describe the development of an iBASFI and to confirm its psychometric properties. This novel methodology included the careful translation of individual BASFI items into instrumented performance-based tests that reflect key activity limitations in patients with axSpA<sup>29,37</sup>.

The concept of “activity limitations” is a relatively new term in the field of SpA and largely replaces the ill-defined “physical function” concept, a key domain in the 1999 ASAS core set to monitor patients with ankylosing spondylitis (AS)<sup>9,11,38</sup>. “Activity limitations” or positively formulated “activity capacity” refers to the difficulty or ability to execute a task in a standardized environment without assistance<sup>10</sup>.

Over the past decades, the assessment of activity limitations in patients with axSpA evolved from self-reported questionnaires<sup>29,39,40,41</sup> toward direct observation/judgment<sup>37</sup> and performance-based tests by an operator. The latter PBM typically quantify movement duration by a hand-held chronometer or repetition counts within a standardized time frame<sup>20,42</sup> reflecting the idea of direct and objective assessment of activity limitation. Our study investigated and innovated these PBM by adding a combination of data from a body-worn, 2-axial accelerometer and fully automated algorithms to detect movement duration. This methodology may have clear advantages over the existing methods.

First, self-reports of activity limitation rely on processes of comprehension, frame of reference (standardized task environment such as chair height), memory retrieval, motivation, emotional status<sup>43</sup>, and response mapping. Although highly feasible and popular in patients with axSpA, they are considered valid at the group level only or to assess change through responder criteria<sup>7</sup>. Similarly, observational methods are also still subject to interpretation bias<sup>44</sup>. Performance-based tests may be less prone to personal and environmental influences<sup>18</sup>, although direct evidence in axSpA is currently not available. Our correlation data with the BASFI largely confirm shared information between techniques (i.e., concurrent construct validity), except for the sock, reach, reach height, and shoulder speed tests. For the

sock test, this mismatch may be explained by variable strategies a patient may use (e.g., put foot on a chair) during daily living, while during the performance-based test, the mode of execution was restricted to sitting for standardization. The latter method ensured close operationalization of the ICF definition of activity capacity (ability in a standardized environment) that contrasts with activity performance (ability in the subject’s real-life environment). For all reaching tasks, the relationship with the BASFI total or corresponding BASFI item is apparently not dependent on arm function, but more on actual body height as seen in the regression analysis. Arm activities may be important for some patients with local shoulder problems, but do not cover the content of the BASFI well.

Second, questionnaires such as the BASFI tend to have floor or ceiling effects and/or just do not pick up longterm changes in activity limitation. For example, several longterm followup studies on anti-TNF now exist<sup>45</sup> and uniformly show a mean BASFI of about 7/10 at baseline which, after initial therapy response, remains stable at about 3/10 during the 8-year followup. In contrast, performance-based tests in AS may be more sensitive to change because they are able to detect improvement even in nonresponders<sup>21</sup> during anti-TNF therapy at 3 months of followup, although learning effects could not be excluded. Similar, in patients scheduled for hip and knee arthroplasty, Stratford, *et al*<sup>46</sup> showed the superiority of performance-based measures to detect increased activity limitation prior to and recovery after surgery in comparison to self-reports. Arguing in favor of our accelerometry-based approach, direct measurement of activities may reveal discriminatory features not included in questionnaires or PBM using chronometers. This was illustrated by the added value of accelerometry in the assessment of turning in patients with Parkinson disease with different stages of disease<sup>47</sup>.

Third, our study found preliminary but overall excellent psychometric properties for all fast-paced and repeated or complex tests. In a study by van Weely, *et al*<sup>20</sup>, 1-week test-retest reliability of chronometer- and performance-based tests ranged between 0.73 and 0.96, indicating a slightly

lower range of reliability compared with our results. The accelerometer used in our study may have limited the variance induced by a physical therapist operating the chronometer. Also, our fully automated feature selection with proven validity in comparison to manual segmentation of signals contrasts with previously reported semiautomated algorithms<sup>24</sup> and favors feasibility. In addition, pilot regression analysis revealed the possibility to obtain a core set of iBASFI items more suitable for clinical practice. The repeated STS movements added most information on activity capacity to the iBASFI. Remarkably, this gross motor task has turned out to be a key-limited, performance-based measure in other disease populations.

There are some limitations to our study. Our proof-of-concept study has a limited sample size that could have affected power. However, in the direct comparison of instrumented performance-based tests with each BASFI item and the total scale to confirm concurrent construct validity, no correlation coefficients of sufficient magnitude to reach the validity criterion showed insignificance. Also, only 4 out of 24 correlation coefficients of insufficient magnitude turned out to be insignificant ( $r$  BASFI scale: reach height, STS;  $r$  BASFI item: sock, lying down). Sample size did not affect regression analysis quality; however, our item reduction is preliminary because cross-validation, responsiveness, and feasibility were not included in the selection procedure<sup>19</sup>. Although we randomly selected a typical sample of outpatients with axSpA, sample size may have affected the generalizability of results. Ongoing research in a larger sample of patients with axSpA will tackle these issues and is needed to prepare the iBASFI for use in clinical practice.

Another limitation may be the reliance on the BASFI to develop the PBM. Although the content validity of this scale is excellent<sup>11</sup>, one is not able to tailor the automated algorithms to patient-specific activities at this point. The preliminary core set of 3 PBM is, however, a good starting point for future research.

Finally, that we did not evaluate all aspects of feasibility of the iBASFI may be a limitation. Performing the extensive protocol took < 20 min on average and we estimate the short iBASFI (height, STS speed test, and cervical rotation) to take < 5 min. This included positioning and loading of the sensor, done by a physical therapist who instructed the patient. Future studies may investigate video instruction and automated detection of sensor location to make the assessment procedure fully automatic<sup>48</sup>. Also, a distinction between a clinical iBASFI and research iBASFI may be considered in future research. An animated activity questionnaire was developed that asked patients with OA to map their activity limitation to videos showing different abilities<sup>49,50</sup>. Because the animated questionnaire correlated highly with PBM (but unexpectedly also self-reports<sup>49</sup>) in these validity experiments, this technique may overcome frame-of-reference issues in self-reports and may reduce the time and

resources inherent to performance-based testing. Future research should compare these techniques head-to-head and elucidate their unique involvement in axSpA outcome assessment.

Our proof-of-concept iBASFI showed promising test-retest reliability, construct validity in comparison with chronometer- and performance-based testing, and construct validity in comparison with the BASFI questionnaire. Future studies should elucidate the added value of these technology-based measures and further validate the iBASFI for use in clinical practice and research.

## ACKNOWLEDGMENT

The authors thank Frederik Adams (physical therapist), all participants, and all staff members who operationally contributed to this study.

## ONLINE SUPPLEMENT

Supplementary data for this article are available online at [jrheum.org](http://jrheum.org).

## REFERENCES

1. Luyten FP, Lories RJ, Verschueren P, de Vlam K, Westhovens R. Contemporary concepts of inflammation, damage and repair in rheumatic diseases. *Best Pract Res Clin Rheumatol* 2006;20:829-48.
2. Sieper J, van der Heijde D, Landewe R, Brandt J, Burgos-Vargas R, Collantes-Estevez E, et al. New criteria for inflammatory back pain in patients with chronic back pain: a real patient exercise by experts from the Assessment of SpondyloArthritis international Society (ASAS). *Ann Rheum Dis* 2009;68:784-8.
3. Elewaut D, Van den Bosch F, Verbruggen G, de Keyser F, Cruyssen BV, Mielants H. Clinical observations programme in SpA: disease parameters, treatment options and practical management issues. *Rheumatol Int* 2009;29:239-50.
4. van der Heijde D, van der Linden S, Bellamy N, Calin A, Dougados M, Khan MA. Which domains should be included in a core set for endpoints in ankylosing spondylitis? Introduction to the ankylosing spondylitis module of OMERACT IV. *J Rheumatol* 1999;26:945-7.
5. van der Heijde D, Calin A, Dougados M, Khan MA, van der Linden S, Bellamy N. Selection of instruments in the core set for DC-ART, SMARD, physical therapy, and clinical record keeping in ankylosing spondylitis. Progress report of the ASAS Working Group. Assessments in Ankylosing Spondylitis. *J Rheumatol* 1999;26:951-4.
6. Sieper J, Rudwaleit M, Baraliakos X, Brandt J, Braun J, Burgos-Vargas R, et al. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. *Ann Rheum Dis* 2009;68 Suppl 2:ii1-44.
7. Anderson JJ, Baron G, van der Heijde D, Felson DT, Dougados M. Ankylosing spondylitis assessment group preliminary definition of short-term improvement in ankylosing spondylitis. *Arthritis Rheum* 2001;44:1876-86.
8. Brandt J, Listing J, Sieper J, Rudwaleit M, van der Heijde D, Braun J. Development and preselection of criteria for short term improvement after anti-TNF alpha treatment in ankylosing spondylitis. *Ann Rheum Dis* 2004;63:1438-44.
9. Boonen A, Braun J, van der Horst Bruinsma IE, Huang F, Maksymowych W, Kostanjsek N, et al. ASAS/WHO ICF Core Sets for ankylosing spondylitis (AS): how to classify the impact of AS on functioning and health. *Ann Rheum Dis* 2010;69:102-7.
10. World Health Organization. International classification of functioning, disability and health. Geneva: World Health Organization; 2001.

11. Sigl T, Cieza A, van der Heijde D, Stucki G. ICF based comparison of disease specific instruments measuring physical functional ability in ankylosing spondylitis. *Ann Rheum Dis* 2005;64:1576-81.
12. Cieza A, Hilfiker R, Boonen A, van der Heijde D, Braun J, Stucki G. Towards an ICF-based clinical measure of functioning in people with ankylosing spondylitis: a methodological exploration. *Disabil Rehabil* 2009;31:528-37.
13. Kiltz U, van der Heijde D, Boonen A, Cieza A, Stucki G, Khan MA, et al. Development of a health index in patients with ankylosing spondylitis (ASAS HI): final result of a global initiative based on the ICF guided by ASAS. *Ann Rheum Dis* 2015;74:830-5.
14. van Weely SF, van Denderen JC, Steultjens MP, van der Leeden M, Nurmohamed MT, Dekker J, et al. Moving instead of asking? Performance-based tests and BASFI-questionnaire measure different aspects of physical function in ankylosing spondylitis. *Arthritis Res Ther* 2012;14:R52.
15. Pedersen OB, Hansen GO, Svendsen AJ, Ejstrup L, Junker P. Adaptation of the Bath measures on disease activity and function in ankylosing spondylitis into Danish. *Scand J Rheumatol* 2007; 36:22-7.
16. Boonen A. Towards a better understanding of the role of psychological variables in arthritis outcome research. *Arthritis Res Ther* 2010;12:106.
17. Brionez TF, Assassi S, Reveille JD, Green C, Leach T, Diekmann L, et al. Psychological correlates of self-reported disease activity in ankylosing spondylitis. *J Rheumatol* 2010;37:829-34.
18. Reneman MF, Jorritsma W, Dijkstra SJ, Dijkstra PU. Relationship between kinesophobia and performance in a functional capacity evaluation. *J Occup Rehabil* 2003;13:277-85.
19. Dobson F, Bennell K, Hinman R, Roos E, Abbott H, Stratford P, et al. OARSI recommended performance-based tests to assess physical function in osteoarthritis of the hip or knee: authors' reply. *Osteoarthritis Cartilage* 2013;21:1625-6.
20. van Weely SF, van Denderen CJ, van der Horst-Bruinsma IE, Nurmohamed MT, Dijkman BA, Dekker J, et al. Reproducibility of performance measures of physical function based on the BASFI, in ankylosing spondylitis. *Rheumatology* 2009;48:1254-60.
21. van Weely SF, van Denderen JC, Steultjens MP, Nurmohamed MT, Dijkman BA, Dekker J, et al. What do we miss? ASAS non-responders on anti-TNF therapy show improvement in performance-based physical function. *Rheumatology* 2013; 52:1884-9.
22. Stratford PW, Kennedy DM. Performance measures were necessary to obtain a complete picture of osteoarthritic patients. *J Clin Epidemiol* 2006;59:160-7.
23. Janssen W, Busmann J, Selles R, Koudstaal P, Ribbers G, Stam H. Recovery of the sit-to-stand movement after stroke: a longitudinal cohort study. *Neurorehabil Neural Repair* 2010;24:763-9.
24. Janssen WG, Busmann JB, Horemans HL, Stam HJ. Validity of accelerometry in assessing the duration of the sit-to-stand movement. *Med Biol Eng Comput* 2008;46:879-87.
25. Janssen WG, Kulcu DG, Horemans HL, Stam HJ, Busmann JB. Sensitivity of accelerometry to assess balance control during sit-to-stand movement. *IEEE Trans Neural Syst Rehabil Eng* 2008;16:479-84.
26. Swinnen TW, Scheers T, Lefevre J, Dankaerts W, Westhovens R, de Vlam K. Physical activity assessment in patients with axial spondyloarthritis compared to healthy controls: a technology-based approach. *PLoS One* 2014;9:e85309.
27. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63:737-45.
28. Osterhaus JT, Purcaru O. Discriminant validity, responsiveness and reliability of the arthritis-specific Work Productivity Survey assessing workplace and household productivity in patients with psoriatic arthritis. *Arthritis Res Ther* 2014;16:R140.
29. Calin A, Garrett S, Whitelock H, Kennedy LG, O'Hea J, Mallorie P, et al. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. *J Rheumatol* 1994;21:2281-5.
30. Van Tubergen A, Debats I, Ryser L, Londoño J, Burgos-Vargas R, Cardiel MH, et al. Use of a numerical rating scale as an answer modality in ankylosing spondylitis-specific questionnaires. *Arthritis Rheum* 2002;47:242-8.
31. Calin A, Jones SD, Garrett SL, Kennedy LG. Bath Ankylosing Spondylitis Functional Index. *Br J Rheumatol* 1995;34:793-4.
32. Haywood KL, M Garratt A, Jordan K, Dziedzic K, Dawes PT. Disease-specific, patient-assessed measures of health outcome in ankylosing spondylitis: reliability, validity and responsiveness. *Rheumatology* 2002;41:1295-302.
33. Jones SD, Calin A, Steiner A. An update on the Bath Ankylosing Spondylitis Disease Activity and Functional Indices (BASDAI, BASFI): excellent Cronbach's alpha scores. *J Rheumatol* 1996;23:407.
34. Pavy S, Brophy S, Calin A. Establishment of the minimum clinically important difference for the bath ankylosing spondylitis indices: a prospective study. *J Rheumatol* 2005;32:80-5.
35. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21:651-7.
36. Landewé R, Dougados M, Mielants H, van der Tempel H, van der Heijde D. Physical function in ankylosing spondylitis is independently determined by both disease activity and radiographic damage of the spine. *Ann Rheum Dis* 2009;68:863-7.
37. Hidding A, van Santen M, De Klerk E, Gielen X, Boers M, Geenen R, et al. Comparison between self-report measures and clinical observations of functional disability in ankylosing spondylitis, rheumatoid arthritis and fibromyalgia. *J Rheumatol* 1994;21:818-23.
38. Boonen A, Rasker JJ, Stucki G. The international classification for functioning, disability and health. A challenge and a need for rheumatology. *Clin Rheumatol* 2007;26:1803-8.
39. Eyres S, Tennant A, Kay L, Waxman R, Helliwell PS. Measuring disability in ankylosing spondylitis: comparison of bath ankylosing spondylitis functional index with revised Leeds Disability Questionnaire. *J Rheumatol* 2002;29:979-86.
40. Dougados M, Gueguen A, Nakache JP, Nguyen M, Amor B. Evaluation of a functional index for patients with ankylosing spondylitis. *J Rheumatol* 1990;17:1254-5.
41. Daltroy LH, Larson MG, Roberts NW, Liang MH. A modification of the Health Assessment Questionnaire for the spondyloarthropathies. *J Rheumatol* 1990;17:946-50.
42. Marcora S, Casanova F, Williams E, Jones J, Elamanchi R, Lemmey A. Preliminary evidence for cachexia in patients with well-established ankylosing spondylitis. *Rheumatology* 2006;45:1385-8.
43. Brionez TF, Assassi S, Reveille JD, Leach TJ, Diekmann L, Ward MM, et al. Psychological correlates of self-reported functional limitation in patients with ankylosing spondylitis. *Arthritis Res Ther* 2009;11:R182.
44. Uhlig T, Lillemo S, Moe RH, Stamm T, Cieza A, Boonen A, et al. Reliability of the ICF Core Set for rheumatoid arthritis. *Ann Rheum Dis* 2007;66:1078-84.
45. Baraliakos X, Haibel H, Fritz C, Listing J, Heldmann F, Braun J, et al. Long-term outcome of patients with active ankylosing spondylitis with etanercept-sustained efficacy and safety after seven years. *Arthritis Res Ther* 2013;15:R67.
46. Stratford PW, Kennedy DM, Riddle DL. New study design



- evaluated the validity of measures to assess change after hip or knee arthroplasty. *J Clin Epidemiol* 2009;62:347-52.
47. King LA, Mancini M, Priest K, Salarian A, Rodrigues-de-Paula F, Horak F. Do clinical scales of balance reflect turning abnormalities in people with Parkinson's disease? *J Neurol Phys Ther* 2012; 36:25-31.
48. Weenk D, van Beijnum BJ, Baten CT, Hermens HJ, Veltink PH. Automatic identification of inertial sensor placement on human body segments during walking. *J Neuroeng Rehabil* 2013;10:31.
49. Peter WF, Loos M, de Vet HC, Boers M, Harlaar J, Roorda LD, et al. Development and preliminary testing of a computerized animated activity questionnaire in patients with hip and knee osteoarthritis. *Arthritis Care Res* 2015;67:32-9.
50. Terwee CB, Coopmans C, Peter WF, Roorda LD, Poolman RW, Scholtes VA, et al. Development and validation of the computer-administered animated activity questionnaire to measure physical functioning of patients with hip or knee osteoarthritis. *Phys Ther* 2014;94:251-61.

**Bath Ankylosing Spondylitis Functional Index Shows Promising Reliability and Validity in the Assessment of Activity Limitations in Axial Spondyloarthritis\***

Swinnen TW, Milosevic M, Van Huffel S, Dankaerts W, Westhovens R, de Vlam K. Bath Ankylosing Spondylitis Functional Index shows promising reliability and validity in the assessment of activity limitations in axial spondyloarthritis. *J Rheumatol* 2016;43:1532-40. The title of this article should be: Instrumented BASFI (iBASFI) shows promising reliability and validity in the assessment of activity limitations in axial spondyloarthritis.

\*This correction is to the First Release version only, published online June 15, 2016.

doi:10.3899/jrheum.150439.C1