Objective Evaluation of Physical Functioning after Tumor Necrosis Factor Inhibitory Therapy in Patients with Ankylosing Spondylitis: A Selection of 3 Feasible Performance-based Tests

Salima F.E. van Weely, Joost Dekker, Martijn P.M. Steultjens, J. Christiaan van Denderen, Michael T. Nurmohamed, Ben A.C. Dijkmans, and Irene E. van der Horst-Bruinsma

ABSTRACT. ***Objective.*** (1) To select a limited number of performance-based tests that are reliable, show improvement in physical functioning after tumor necrosis factor inhibitor (TNFi) therapy in patients with ankylosing spondylitis (AS), and generate information equivalent to the full set of tests, and (2) are feasible for use in daily clinical practice.

Methods. Eight performance-based tests were evaluated. To eliminate redundant testing, the tests that showed adequate reliability, the highest standardized response mean (SRM), and the largest proportion of patients with an improved performance-based physical functioning were selected. The selected tests were combined into a new criterion for improvement in physical functioning (AS Performance-based Improvement; ASPI). The number and percentage of improved patients identified with the ASPI and identified with the full set of performance tests were compared.

Results. Reliability for all tests was adequate to excellent (ICC 0.73–0.96). The tests for bending, putting on socks, and getting up from the floor had the highest SRM (0.52–0.74) and showed the largest proportion of improved patients after TNFi therapy. The combination of these 3 tests was feasible in daily clinical practice and showed improved physical functioning after TNFi therapy in 67% of the patients.

Conclusion. The 3 selected tests are recommended for use in daily practice because they generate information comparable to the full set. They are reliable and feasible, and the combination of these tests showed improved physical functioning after TNFi therapy in 67% of the patients. Evaluation of physical functioning might be improved by adding these tests to other AS outcome measures. (First Release Jan 15 2015; J Rheumatol 2015;42:623–9; doi:10.3899/jrheum.140337)

Key Indexing Terms:
ANKYLOSING SPONDYLITIS OUTCOME ASSESSMENT FEASIBILITY STUDY
PHYSICAL FUNCTION BASFI

From Reade, Centre for Rehabilitation and Rheumatology; Department of Rehabilitation Medicine, and the Department of Rheumatology, and the Department of Psychiatry, EMGO Institute, Vrije Universiteit (VU) University Medical Centre, Amsterdam, the Netherlands; and the Institute for Applied Health Research and School of Health and Life Sciences, Glasgow Caledonian University, Glasgow, Scotland, UK.

Supported by an unrestricted financial grant from the Dutch Arthritis Foundation.

S.F.E. van Weely, MSc; J.C. van Denderen, MD; M.T. Nurmohamed, MD, PhD, Reade, Centre for Rehabilitation and Rheumatology; J. Dekker, PhD, Reade, Centre for Rehabilitation and Rheumatology, and the Department of Rehabilitation Medicine, Department of Psychiatry, EMGO Institute, VU University Medical Centre; B.A.C. Dijkmans, MD, PhD, Reade, Centre for Rehabilitation and Rheumatology, and the Department of Rheumatology, VU University Medical Centre; I.E. van der Horst-Bruinsma, MD, PhD, Department of Rheumatology, VU University Medical Centre; M.P.M. Steultjens, PhD, Institute for Applied Health Research and School of Health and Life Sciences, Glasgow Caledonian University.

Address correspondence to S.F.E. van Weely, Reade, Centre for Rehabilitation and Rheumatology, Jan van Breemenstraat 2, P.O. Box 58271, 1040 HG Amsterdam, the Netherlands. E-mail: s.v.weely@reade.nl

Accepted for publication December 2, 2014.

Ankylosing spondylitis (AS) is characterized by inflammation, pain, stiffness, and fusion of the spine, leading to limitations in physical functioning, work disability, and impaired quality of life[1]. Maintaining or improving physical functioning is one of the main treatment goals. Physical functioning is also considered an important outcome measure for the evaluation of the disease course and effectiveness of therapy[2,3].

For assessing physical functioning, the patient-reported, disease-specific, reliable, and responsive Bath AS Functional Index (BASFI) is used most commonly[4,5,6,7]. However, physical functioning is not a single variable but rather a collection of different health concepts that together paint a picture of how a disease affects a patient in daily life[8]. Self-reported measures such as the BASFI can be influenced by discrepancies between perceptions of a person's ability and the patient's actual performance (underestimation or overestimation), and therefore only show perceived limita-

tions in physical functioning rather than true limitations. Such discrepancies can occur as a result of personality traits, depression, poor cognitive function, language, educational level, expectations, and pain[9,10,11,12,13,14,15,16,17].

In an effort to identify more objective outcome measures that assess actual limitations in physical functioning in AS, 8 performance-based tests based on items of the BASFI were developed[18]. In these tests, the actual time patients needed to complete a task was measured. The tests showed to be reliable[18] and provided more objective information on physical functioning because patients with AS seemed to incorporate exertion and pain in their assessment of perceived physical functioning on the BASFI[19]. Further, an improvement in physical functioning after 3 months of tumor necrosis factor inhibitors (TNFi) treatment was shown in patients with AS who were classified as nonresponders according to the Assessment of SpondyloArthritis international Society 20% improvement criteria (ASAS20)[20,21,22]. Thus, in patients with AS, performance-based tests can provide a better estimation of a person's abilities. However, performing all 8 tests might be too strenuous for patients and cumbersome in daily clinical practice. In research and clinical practice, it is important to eliminate redundant testing to save energy, time, and money. Therefore, our study aimed to (1) select a limited number of performance-based tests that are reliable, show improvement in physical functioning after TNFi therapy, and generate information equivalent to the full set, and (2) are feasible for use in daily clinical practice.

MATERIALS AND METHODS

From May 2006 to June 2010, adult patients (≥ 18 yrs) fulfilling the modified New York criteria[23] and having sufficient command of the Dutch language were consecutively recruited from a large outpatient center for rheumatology and rehabilitation, Reade in Amsterdam. Patients were excluded if they had pulmonary, cardiovascular, or neurological comorbidity affecting their ability to perform daily activities. The local medical ethical committee approved the study and all patients gave written informed consent.

Performance-based tests of physical functioning. The 8 performance-based tests used in our study were based on items of the BASFI and consisted of (1) climbing stairs, (2) bending (picking up pens from the floor), (3) reaching (putting pens on a high shelf), (4) putting on socks, (5) rising up and sitting down on a chair, (6) getting up from the floor, (7) looking over the shoulder, and (8) a physically demanding activity (shuttle-walk test). The tests were executed as described by van Weely, *et al*[18]. Patients were instructed to perform the test at their own pace, though as quickly as possible and to stop if necessary. Outcome of the performance tests was the time needed to complete the task (in seconds), except for 7 (looking over the shoulder), in which the range of vision was recorded in points. Tests 4–6 (putting on socks, rising up and sitting down on a chair, and getting up from the floor) were performed 3 times and the mean performance times were used.

Study design. The flow chart in Figure 1 illustrates the inclusion of patients. All patients were assessed at baseline. A test-retest design with a 1-week interval was used to evaluate the reliability of the performance-based tests. A longitudinal design was used to evaluate the improvement in physical functioning after TNFi therapy. Patients eligible for TNFi therapy were reassessed after 3 months of TNFi treatment. The assessments included completion of the 8 performance-based tests and the BASFI.

For the reliability analysis, we used a summary of previous results[18]. The results of improvements in performance-based physical functioning after TNFi therapy were based on the secondary analyses of a previous study[22].

Statistical analyses. All analyses were performed by using SPSS for Windows 18.0 (SPSS Inc.). Patient characteristics and scores of the performance-based tests before and after TNFi treatment were examined by calculating means and SD for all continuous data and percentages for categorical data.

Reliability was assessed by calculating ICC (ICC 2.1.A). An ICC is an adequate measure of reliability and has the capability of differentiating among patients, taking into account both systematic errors between 2 measurements and random measurement error. An ICC of > 0.70 is required for the comparison of groups, whereas an ICC > 0.90 is recommended for individual evaluation[24,25].

Standardized response means (SRM, mean change ÷ SD change) were calculated to provide information about the importance of the identified differences between the tests. A higher SRM indicates a greater effect or clinically important change. SRM of 0.2, 0.5, and 0.8 or above represents small, moderate, and large clinical changes, respectively[26].

Improvement in performance-based physical functioning after TNFi therapy was defined by an intraindividual improvement of ≥ 20%. This cutoff point was chosen analogous to the ASAS20 improvement criteria[20,21] and may be considered a clinically meaningful improvement. As opposed to absolute changes, relative, intraindividual changes provided a better reflection of clinically meaningful changes for individual patients. This is a commonly used approach for defining improvement in rheumatology for both research and clinical practice. For each test, the proportion (number and percentage) of patients with an improved performance-based physical functioning was calculated.

Selection of test for the AS Performance-based Improvement (ASPI) criterion. The performance-based tests with the highest reliability, highest SRM, and largest proportion of patients with an improved performance-based physical functioning were selected, and possible issues concerning the practical applicability of the selected tests were considered. The selected tests were combined into a new criterion (i.e., the ASPI). This criterion was composed of an intraindividual improvement in physical functioning of ≥ 20% after TNFi therapy on 1 or more of the selected performance-based tests and the absence of deterioration on the potential remaining tests. Deterioration in physical functioning was defined as a worsening of ≥ 20%. Again, the cutoff value of an intraindividual improvement of ≥ 20% was chosen analogous to the ASAS20 improvement criteria[20,21] and may be considered to reflect a clinically meaningful improvement. A definition of improvement that took multiple domains (i.e., tests) into account was chosen because it had greater content validity, because consistent improvement had to be present and deterioration had to be absent[21]. The number and percentage of improved patients identified with the ASPI and identified with the full set of tests were compared.

RESULTS

Patient characteristics. Table 1 displays the patient characteristics of 124 patients (69% men) with a mean age (± SD) of 46.0 years (11.6) and disease duration (± SD) of 14.5 (10.2). The majority of patients was HLA-B27–positive (86%) and used nonsteroidal antiinflammatory drugs (67%). The most frequently occurring extraspinal symptoms were peripheral arthritis (42%) and uveitis (36%). All patients fulfilled the eligibility criteria and none of the patients were excluded because of pulmonary, cardiovascular, or neurological comorbidity affecting the patients' ability to perform daily activities.

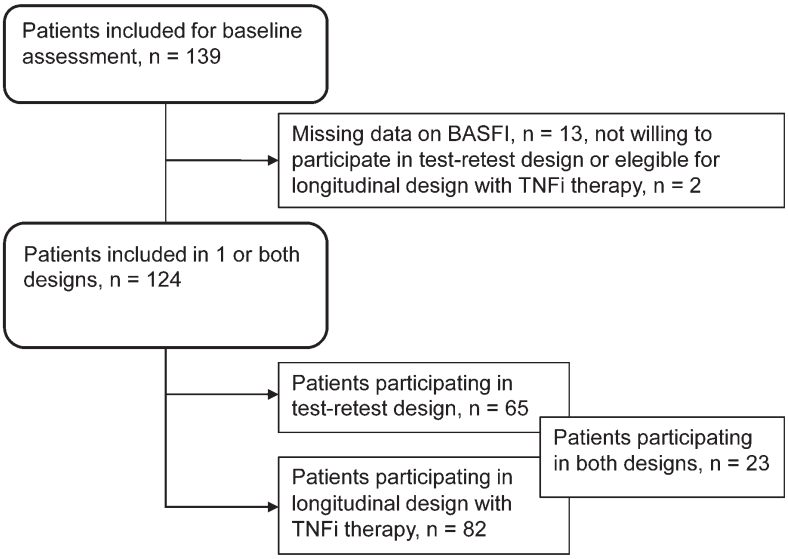

Figure 1. Patient flow chart. BASFI: Bath Ankylosing Spondylitis Functional Index; TNFi: tumor necrosis factor inhibitor.

Table 1. Baseline characteristics of all patients (n = 124). Values are mean ± SD or n (%).

Characteristic	Value
Men	85 (69)
Age, yrs	46.0 ± 11.6
Symptom duration, yrs	21.8 ± 12.1
Disease duration, yrs	14.5 ± 10.2
Medication	
None	19 (16)
NSAID	81 (67)
Biologicals	4 (3)
DMARD*/combination	17 (14)
HLA-B27+	103 (86)
ESR, mm/h	21.6 ± 19.3
Extraspinal symptoms	
Psoriasis	8 (7)
Uveitis	44 (36)
Inflammatory bowel disease	6 (5)
Arthritis	52 (42)
BASFI, score 0–10	4.9 ± 2.3
BASDAI, score 0–10	5.2 ± 2.3
BASMI, score 0–10	4.1 ± 1.8

* Sulfasalazine, methotrexate. NSAID: nonsteroidal antiinflammatory drugs; DMARD: disease-modifying antirheumatic drugs; ESR: erythrocyte sedimentation rate; BASFI: Bath Ankylosing Spondylitis Functional Index; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASMI: Bath Ankylosing Spondylitis Metrology Index.

For the reliability analyses, 65 consecutive patients were included. Of the 65 patients included in the test-retest design, 42 patients participated only in this design and not in the longitudinal design. Of these 42 patients, 4 (3.2% of total population) were already stable on therapy with biologicals. Improvement in physical functioning after TNFi therapy was evaluated in 82 patients who were eligible for TNFi treatment (Figure 1). Fifty-seven patients were treated with etanercept and 25 patients were treated with adalimumab.

Characteristics of performance-based tests. For all tests, aids and instructions for the patients were necessary. In 4 tests (bending, putting on socks, rising up and sitting down on a chair, and getting up from the floor), only simple aids (i.e., chair, mat) were necessary. Other tests required more complex aids (i.e., stairs, horizontal board with numbers and characters, walking space, and heart rate monitoring device). Most tests only required simple instructions (i.e., climbing stairs, bending, reaching, putting on socks, rising up and sitting down on a chair, and getting up from the floor) whereas 2 tests (looking over the shoulder and performing the shuttle-walk test) needed more complex instructions. All tests were easy to conduct in small environments, except for climbing stairs and the shuttle-walk test. The tests were easy to administer and no practice trials were necessary.

The tests were well tolerated; almost all patients (95%) were able to perform all 8 tests, and only 4 patients had trouble with the test for putting on socks. Table 2 shows the duration for each test in seconds (except for test 7, looking over the shoulder, which is measured in points) for all 124 patients and the 82 who were eligible for TNFi therapy. The shortest test times were recorded for climbing stairs and getting up from the floor, and the longest time was seen in doing a physically demanding activity (i.e., the shuttle-walk test).

Table 2. Characteristics of performance-based tests of all patients (n = 124) and patients eligible for TNFi therapy (n = 82) at baseline and after 3 months of treatment and SRM. Values are mean test time in seconds ± SD unless otherwise specified.

Performance-based Tests	All Patients, Baseline, n = 124	Before TNFi Therapy, Baseline, n = 82	After 3 Mos of TNFi Therapy, n = 82	SRM, n = 82
1. Climbing stairs	6.0 ± 2.8	6.1 ± 3.3	5.4 ± 1.8[c]	0.31
2. Bending	19.5 ± 10.3	19.6 ± 10.0	16.3 ± 10.1[d]	0.71
3. Reaching	11.1 ± 3.3	11.3 ± 3.7	10.2 ± 2.6[d]	0.42
4. Putting on socks	19.0 ± 11.0	18.4 ± 10.1	15.4 ± 8.9[d]	0.60
5. Rising up and sitting down on a chair	12.1 ± 6.7	12.4 ± 6.9	10.2 ± 4.2[d]	0.46
6. Getting up from the floor	8.4 ± 6.2	8.7 ± 6.9	6.5 ± 3.9[d]	0.50
7. Looking over the shoulder[a,b]	23.0 ± 6.4	23.2 ± 6.6	24.2 ± 6.0[c]	0.40
8. Physically demanding activity[b]	444.0 ± 114.8	447.1 ± 124.0	477.6 ± 111.0[d]	0.46

[a] Test 7 in points. [b] Higher score is related to a better functional performance. [c] p value < 0.01. [d] p value < 0.001 for the difference in test time before and after TNFi therapy (n = 82). TNFi: tumor necrosis factor inhibitors; SRM: standardized response means.

Reliability of the performance-based tests. All tests had a good to excellent reliability[18]. The ICC varied between 0.73 and 0.96. The highest levels of test-retest reliability were found for tests 4, 5, 7, and 8 (putting on socks, rising up and sitting down on a chair, looking over the shoulder, and physically demanding activity), with ICC ranging from 0.90 to 0.96. An ICC of > 0.90 was high enough for evaluations on an individual level. Only slightly lower but still adequate ICC (> 0.70) were seen in tests 1, 2, 3, and 6 (climbing stairs, bending, reaching, and getting up from the floor).

Improvement in performance-based physical functioning after TNFi therapy. After TNFi therapy, performance-based physical functioning improved, which was illustrated by a decrease in test time for tests 1–6 and an increased score and time for tests 7 and 8, respectively (all p values < 0.05; Table 2). Thus, patients were able to perform activities more quickly (tests 1–6), could look farther over their shoulder (test 7), and could endure a physically demanding activity longer (test 8).

Table 2 also shows the SRM. Tests 2, 4, and 6 (bending, putting on socks, getting up from the floor) showed an SRM of 0.50 or higher, indicating a moderately important clinical improvement. The other tests had a lower SRM.

In Figure 2, for each test, the number and proportion of patients with an intraindividual improvement of ≥ 20% in physical functioning after 3 months of TNFi therapy are

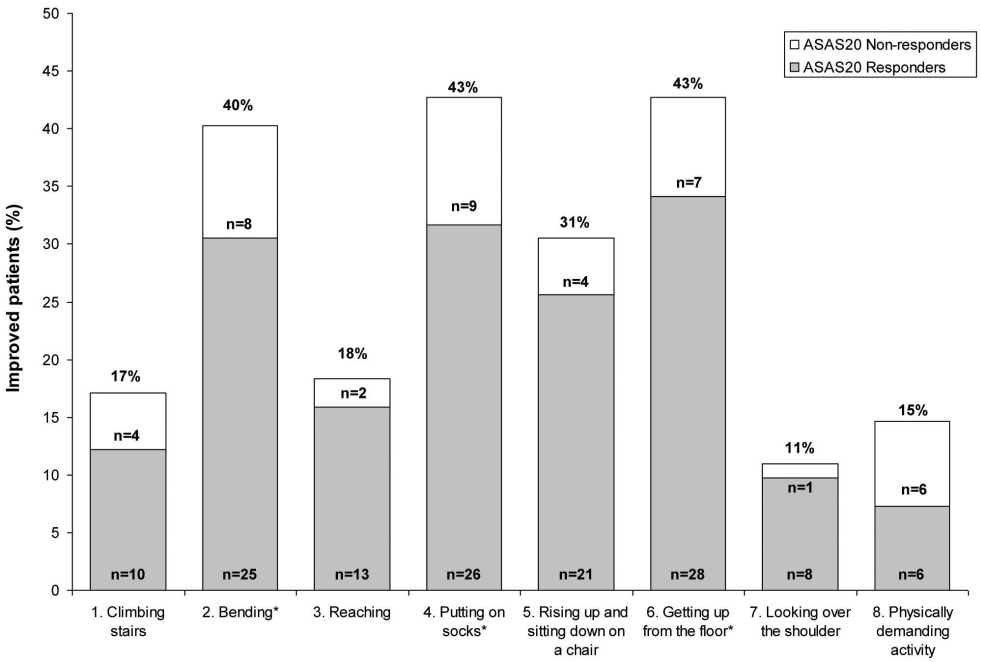

Figure 2. Number and percentage of patients with improved physical functioning after 3 months of TNFi treatment (n = 82). * Tests 2, 4, and 6 (bending, putting on socks, and getting up from the floor) were selected and combined into a new criterion for performance-based improvement (ASPI). TNFi: tumor necrosis factor inhibitors; ASPI: Ankylosing Spondylitis Performance-based Improvement; ASAS20: Assessment of SpondyloArthritis international Society 20% improvement criteria.

shown. The largest proportions of improved patients were seen in tests 2, 4, 5, and 6 (bending, putting on socks, rising up and sitting down on a chair, and getting up from the floor); 40%, 43%, 31%, and 43%, respectively. For the other tests, the percentages of improvers were lower and ranged from 11% to 18%.

In Figure 2, a distinction is made between the number of physically improved patients in ASAS20 responders and nonresponders. It shows that in tests 1, 2, 4, 6, and 8 (stair climbing, bending, putting on socks, getting up from the floor, and the shuttle-walk test), 20% or more of the physically improved patients were classified as ASAS20 nonresponders to TNFi therapy.

Selection of performance-based tests. All tests had an adequate reliability; therefore, none of the performance tests was rejected based on these results. Based on the results of the SRM and improvement in performance-based physical functioning after TNFi therapy, tests 2, 4, and 6 (bending, putting on socks, and getting up from the floor) were selected. These tests showed an SRM of > 0.50 and the largest proportion of improved patients (all tests > 40%). Test 5 (rising up and sitting down on a chair) did not reach these thresholds and was rejected based on this result. Tests 2, 4, and 6 (bending, putting on socks, and getting up from the floor) were easily applicable in daily clinical practice. Despite 4 patients having trouble with test 4 (putting on socks), this test was nevertheless selected because of the high reliability and responsiveness.

Combination of performance test in ASPI. The 3 selected tests (bending, putting on socks, and getting up from the floor) were combined into a new criterion for improvement in performance-based physical functioning (i.e., the ASPI). Patients were defined as improvers on the ASPI if they had an intraindividual improvement in physical functioning of ≥ 20% on at least 1 of the 3 selected performance-tests and an absence of deterioration on the potential remaining tests. Deterioration was defined as worsening of ≥ 20%. Thus, the final ASPI was referring to whether a patient was or was not classified as an improver based on improvement (in seconds) on 1 or more of the 3 selected tests and absence of deterioration on the remaining test(s).

After 3 months of TNFi therapy, the majority of patients (67%, n = 55/82) was classified as improvers according to the ASPI criterion. This equaled the percentage of patients identified with the full set of tests (63%, n = 52/82). Also the percentage of improvers in ASAS20 nonresponders found (18%, n = 15/82) was comparable to the full set (16%, n = 13/82; Table 3).

Table 3. Cross-tabulation between improvement in performance-based physical functioning (ASPI)[a] and the ASAS20 response after 3 months of TNFi therapy (n = 82). Values are n.

Performance-based Physical Functioning, ASPI[a]	ASAS20 Response		
	Nonresponder	Responder	Total
Non-improver	12	15	27
Improver	15	40	55
Total	27	55	82

[a] Intraindividual improvement of ≥ 20% on at least 1 test (bending, putting on socks, getting up from the floor) and absence of deterioration in performance-based physical functioning on the potential remaining tests. Deterioration defined as worsening of ≥ 20%. ASPI: Ankylosing Spondylitis Performance-based Improvement; ASAS20: Assessment of SpondyloArthritis international Society 20% improvement criteria; TNFi: tumor necrosis factor inhibitors.

DISCUSSION

The ASPI, a combination of 3 reliable performance-based tests, can show improvement in physical functioning after TNFi treatment and generate information equivalent to the full set of tests. The 3 standardized tests (bending, putting on socks, and getting up from the floor) are easy to administer, well-tolerated by patients with varying limitations in physical functioning, and are feasible in daily clinical practice.

The ASPI showed an improvement in physical functioning after TNFi therapy in 67% of the patients with AS. This is comparable to the number of patients who improved according to the ASAS20 response criteria in our study and other published studies[27,28]. However, 18% of the patients showed improvement in performance-based physical functioning but were ASAS20 nonresponders. By performing the ASPI, more objective information on improvement in physical functioning is obtained and a better estimation of the ability of patients with AS can be provided.

The time and equipment needed to obtain information on improvement in physical functioning are limited; a stopwatch, a shelf or table, 6 pens, a pair of socks, a chair, a mat, and a maximum of 15 min to perform the 3 activities are necessary. One instruction session was sufficient to warrant a good execution of these tests by a nurse or a physical or occupational therapist. Thus, little effort and material are necessary to perform the 3 tests (bending, putting on socks, and getting up from the floor) and to identify an additional 18% of patients with improvements in physical functioning after TNFi therapy. These results provide support for the use of 3 performance-based tests in daily clinical practice, in addition to, for example, the ASAS20 response criteria.

To our knowledge, ours is the first study that provides a feasible and objective outcome measure to evaluate the domain physical functioning in patients with AS after TNFi therapy. Until recently, interventions in AS were merely evaluated using patient-reported outcome measures. However, contrary to performance-based tests, questionnaires can be influenced by overestimation or underestimation by

the patient. Therefore, there is a need for more objective outcome measures in the evaluation of therapies in AS. For the domain disease activity, the AS Disease Activity Score (ASDAS)[29] has been developed. In this outcome measure, an objective serum variable [C-reactive protein (CRP) or erythrocyte sedimentation rate] is included next to self-reported questions on symptoms of disease activity. However, a disadvantage of the ASDAS is that acute-phase reactants such as CRP are not always raised in patients with AS who have high disease activity. Hence, for an objective assessment in patients with AS, performance-based tests could have an important additional value in the assessment of efficacy of treatment.

A cutoff value to define improvement on the performance-based tests of physical functioning is not available. In our study, an intraindividual improvement of ≥ 20% was used, analogous to the ASAS20 response criteria. A publication by Tubach, *et al* supported the use of this cutoff value[30]. In their prospective multinational study, an estimation of the minimum clinically important improvement (MCII) across various diseases (e.g., AS and chronic back pain), countries, and outcomes was made. They promoted the use of 20% relative improvement as a value for the MCII[30]. In future, a value for absolute improvement in performance-based physical functioning and perhaps an alternative scoring system has to be determined to improve the suitability of these tests in clinical trials.

Our study provides important information on the value and use of performance-based tests in patients with AS. However, the selected study population cannot be characterized as a common AS population because of the relatively long disease duration (14.5 ± 10.2) and high percentage of (patient-reported) uveitis and arthritis (36% and 42%, respectively). Therefore, replication of our study in larger and different study populations and after other interventions than TNFi therapy is necessary. Future research should also give more insight into which factors are underlying limitations in performance-based or self-reported physical functioning, to better comprehend this complex domain. Also, repeating the measures after a longer followup duration is necessary to further evaluate the value of the performance-based tests in identifying changes in physical functioning in the longterm.

Our study shows that the tests for bending, putting on socks, and getting up from the floor are easily applicable and feasible for use in daily clinical practice. A new criterion, the ASPI, using a combination of these 3 tests, showed an improved physical functioning after TNFi therapy in 67% of the patients, of which 18% were ASAS20 nonresponders. Information on improvements after TNFi therapy, in addition to self-reported outcome measures such as the ASAS20 response, is easily gained. By performing only 3 instead of all performance-based tests, redundant testing is eliminated and time, money, and energy can be saved. In future, evaluations of the effectiveness of TNFi therapy in patients with AS might be improved by adding these tests to other outcome measures.

REFERENCES

1. Braun J, Sieper J. Ankylosing spondylitis. Lancet 2007; 369:1379-90.
2. 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.
3. van der Heijde D, Bellamy N, Calin A, Dougados M, Khan MA, van der Linden S. Preliminary core sets for endpoints in ankylosing spondylitis. Assessments in Ankylosing Spondylitis Working Group. J Rheumatol 1997;24:2225-9.
4. 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.
5. Auleley GR, Benbouazza K, Spoorenberg A, Collantes E, Hajjaj-Hassouni N, van der Heijde D, et al. Evaluation of the smallest detectable difference in outcome or process variables in ankylosing spondylitis. Arthritis Rheum 2002;47:582-7.
6. Calin A, Nakache JP, Gueguen A, Zeidler H, Mielants H, Dougados M. Outcome variables in ankylosing spondylitis: evaluation of their relevance and discriminant capacity. J Rheumatol 1999;26:975-9.
7. Ruof J, Stucki G. Comparison of the Dougados Functional Index and the Bath Ankylosing Spondylitis Functional Index. A literature review. J Rheumatol 1999;26:955-60.
8. Zochling J, Braun J, van der Heijde D. Assessments in ankylosing spondylitis. Best Pract Res Clin Rheumatol 2006;20:521-37.
9. Kivinen P, Sulkava R, Halonen P, Nissinen A. Self-reported and performance-based functional status and associated factors among elderly men: the Finnish cohorts of the Seven Countries Study. J Clin Epidemiol 1998;51:1243-52.
10. Guralnik JM, Branch LG, Cummings SR, Curb JD. Physical performance measures in aging research. J Gerontol 1989;44:M141-6.
11. Hoeymans N, Feskens EJ, van den Bos GA, Kromhout D. Measuring functional status: cross-sectional and longitudinal associations between performance and self-report (Zutphen Elderly Study 1990-1993). J Clin Epidemiol 1996;49:1103-10.
12. Myers AM, Holliday PJ, Harvey KA, Hutchinson KS. Functional performance measures: are they superior to self-assessments? J Gerontol 1993;48:M196-206.
13. Kempen GI, van Heuvelen MJ, van den Brink RH, Kooijman AC, Klein M, Houx PJ, et al. Factors affecting contrasting results between self-reported and performance-based levels of physical limitation. Age Ageing 1996;25:458-64.
14. Terwee CB, van der Slikke RM, van Lummel RC, Benink RJ, Meijers WG, de Vet HC. Self-reported physical functioning was more influenced by pain than performance-based physical functioning in knee-osteoarthritis patients. J Clin Epidemiol 2006;59:724-31.
15. Kennedy D, Stratford PW, Pagura SM, Walsh M, Woodhouse LJ. Comparison of gender and group differences in self-report and physical performance measures in total hip and knee arthroplasty candidates. J Arthroplasty 2002;17:70-7.
16. Wittink H, Rogers W, Sukiennik A, Carr DB. Physical functioning: self-report and performance measures are related but distinct. Spine 2003;28:2407-13.
17. Wand BM, Chiffelle LA, O'Connell NE, McAuley JH, Desouza LH. Self-reported assessment of disability and performance-based

assessment of disability are influenced by different patient characteristics in acute low back pain. Eur Spine J 2010;19:633-40.
18. van Weely SF, van Denderen CJ, van der Horst-Bruinsma IE, Nurmohamed MT, Dijkmans BA, Dekker J, et al. Reliability of performance measures of physical function based on the BASFI, in ankylosing spondylitis. Rheumatology 2009;48:1254-60.
19. 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.
20. 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.
21. 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.
22. van Weely SF, van Denderen JC, Steultjens MP, Nurmohamed MT, Dijkmans 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.
23. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361-8.
24. de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. J Clin Epidemiol 2006;59:1033-9.
25. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use, 4th ed. New York: Oxford University Press; 2003.
26. Garratt AM, Ruta DA, Abdalla MI, Russell IT. SF 36 health survey questionnaire: II. Responsiveness to changes in health status in four common clinical conditions. Qual Health Care 1994;3:186-92.
27. Dijkmans B, Emery P, Hakala M, Leirisalo-Repo M, Mola EM, Paolozzi L, et al. Etanercept in the longterm treatment of patients with ankylosing spondylitis. J Rheumatol 2009;36:1256-64.
28. Braun J, Deodhar A, Dijkmans B, Geusens P, Sieper J, Williamson P, et al; Ankylosing Spondylitis Study for the Evaluation of Recombinant Infliximab Therapy Study Group. Efficacy and safety of infliximab in patients with ankylosing spondylitis over a two-year period. Arthritis Rheum 2008;59:1270-8.
29. Lucas C, Landewé R, Sieper J, Dougados M, Davis J, Braun J, et al. Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68:18-24.
30. Tubach F, Ravaud P, Martin-Mola E, Awada H, Bellamy N, Bombardier C, et al. Minimum clinically important improvement and patient acceptable symptom state in pain and function in rheumatoid arthritis, ankylosing spondylitis, chronic low back pain, hand osteoarthritis, and hip and knee osteoarthritis: Results from a prospective multinational study. Arthritis Care Res 2012; 64:1699-707.