

VU Research Portal

Physical functioning in Ankylosing Spondylitis patients

van Weelij, S.F.E.

2015

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

van Weelij, S. F. E. (2015). *Physical functioning in Ankylosing Spondylitis patients: Performance-based assessment and prediction*. [PhD-Thesis – Research external, graduation internal, Vrije Universiteit Amsterdam].

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Chapter 7

General discussion

The aim of the research described in this thesis was to develop performance-based tests of physical functioning in patients with Ankylosing Spondylitis (AS) and to gain more insight into the course and predictors of physical functioning in AS patients treated with tumour necrosis factor alfa inhibitors (TNFi). In this final chapter, the main results of the studies will be summarised and discussed. Furthermore, suggestions for future research and clinical implications will be given.

Development of performance-based tests of physical functioning

As limitations in physical functioning are an important feature of AS, adequate measurement of physical functioning is essential. Commonly, limitations in physical functioning are assessed with the self-reported BASFI questionnaire [1]. However, questionnaires report what a patient thinks he/she can do, which can be an overestimation or underestimation of the actual level of functioning. Hence, there is an unmet need for a new outcome parameter that is less influenced by patient perceptions. Performance-based measures could fulfil this need. These tests are frequently used in other diseases and conditions (e.g., the elderly, and those with chronic low back pain, respiratory disease, heart failure, and osteoarthritis [2-8]), to measure limitations in physical functioning, but were not yet available in AS. We therefore developed eight performance-based measures based on the BASFI questionnaire: 1. climbing stairs, 2. bending, 3. reaching, 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) (Appendix III). These tests could offer a valid way to assess changes in physical functioning over time and evaluate the response to therapeutic interventions like TNFi therapy.

The quality of an assessment tool like performance-based tests depends on its measurement properties, such as reliability (the degree to which the measurement is free from measurement error), validity (the degree to which an instrument measures the construct it purports to measure), and responsiveness (the ability of a measurement instrument to detect change over time in the construct to be measured) [9]. The feasibility of a test is also an important feature. These properties are (highly) dependent on the population in which they are examined. Thus, before an outcome measure like the performance-based tests can be used in research or clinical practice, it is important to evaluate the measurement properties in the intended patient population [9].

In **Chapter 2**, the intra-rater test-retest reliability and measurement error – both important aspects of reliability – of the eight performance-based tests are reported. Test-retest reliability refers to the extent to which scores are the same over time for patients who have not changed [9]. This was assessed with intraclass correlation coefficients (ICC). Measurement error refers

to the systematic and random error of a patient's score that is not attributable to true changes in the construct to be measured [9]. The random error was estimated by calculating the standard error of measurement (SEM) and minimal detectable difference (MDD). The method of limits of agreement (LoA) was used to assess the systematic error.

All the aforementioned parameters were reported for both the performance scores (i.e.: time) and the scores for exertion and pain associated with executing the tests. Patients were examined twice by the same observer, with a 1-week interval. During this short period, it was assumed that the disease activity remained stable and the level of physical functioning was unchanged. The observed differences would therefore be attributable to measurement error. The reliability of the performance score (ICC) was adequate to excellent for all tests and satisfactory for group assessment (research context) (i.e., all ICC >0.70). For the evaluation of individual patients, a higher level of reliability of the performance score is desirable for some tests (i.e., 'bending' and 'reaching'). The reliability of the exertion and pain scores associated with the execution of the tests also showed adequate ICC, with the exception of the test for 'looking over the shoulder'. The ICC for exertion and pain associated with this test were just below the threshold value of 0.70.

The ICC of the performance scores were comparable with the ICC values reported for the reliability of the BASFI questionnaire. There was no systematic difference between the first and second measurement, with the exception of the performance score for the 'bending' test. The MDD ranged between 10%-24% of the observed score. To increase the reliability and decrease the SEM and MDD, the performance-based tests should consist of repeated measurements.

Another aspect of reliability is the internal consistency (the degree of the interrelatedness among the items [9]) that was reported in **Chapter 3**. It was shown that all tests measure the same construct, except for the test 'looking over the shoulder'. This test had low inter-item correlations with all other tests (0.25-0.32), which was confirmed by factor analyses and Cronbach's Alpha. Therefore, later analyses were performed separately for the 'looking over the shoulder' test and the combination of the remaining seven tests.

In order to assess the validity of the performance-based tests, cross-sectional associations between the performance tests and the BASFI questionnaire were examined and reported in **Chapters 2 and 3**. Strong correlations suggest that both outcomes are measuring the same construct, whereas weak correlations suggest a small overlap between the two methods, which apparently are measuring different aspects of the same construct. Patients completed the BASFI questionnaire and executed the performance-based tests within a 1-month period. The assessor of the performance tests was blinded for the results of the BASFI questionnaire. The univariable associations between the performance scores (measured in time) and BASFI were moderate to weak. This was seen both on item level (e.g., the test for 'stair climbing'

with the corresponding BASFI question for ‘stair climbing’) as well as on the total score of the performance tests with the total BASFI score. Multivariable analyses showed that exertion and pain associated with the execution of the performance tests related stronger to the BASFI questionnaire than actual performance (i.e., time). The weak associations between the performance tests and questionnaire found in our study are in line with other studies [7, 8, 10-15], and suggest that performance-based tests of physical functioning and the BASFI questionnaire do not measure the same aspects of physical functioning in patients with AS. Furthermore, the performance-based test showed to have a moderate association with the BASMI and a weak association with the BASDAI questionnaire. This suggests that actual physical functioning is more influenced by spinal and hip mobility (BASMI) than by disease activity (BASDAI).

In **Chapter 4** the responsiveness (the ability to detect change over time in the construct to be measured) was assessed. To determine whether the performance-based tests were sensitive enough to detect change over time, the tests were applied to a group of AS patients eligible for TNFi therapy. TNFi therapy is known to be a very effective treatment of AS. Patients were examined at baseline and after 3 months of treatment.

Improvement in performance-based physical functioning was defined as a 20% intra-individual improvement from baseline. This method is commonly used in rheumatology and proposed in literature [16]. The Assessment of Spondyloarthritis international Society (ASAS) also adopted an intraindividual change in terms of percentage for defining improvement as a basic principle and this resulted – among others – in the ASAS20% improvement criteria [17]. The advantage of such a relative, intraindividual change is that is easier to interpret in a clinical setting.

After 3 months of TNFi treatment, a significant improvement in performance-based physical functioning was seen ($p < 0.0001$) and 63% of the patients were defined as improvers (i.e., showed an improvement of $\geq 20\%$). Almost 50% of the patients had an improvement in performance-based functioning *and* were classified as ‘responders’ according to the ASAS20 criteria. A substantial proportion of patients, however, improved on the performance tests but were not responders on the ASAS20. Patients thus improved in performance (e.g., could get up from the floor more quickly), but (still) experienced pain, stiffness, fatigue, and/or unchanged limitations in physical functioning. The same pattern was seen when improvers on the performance tests were compared to improvers on the BASFI questionnaire. Most patients improved on both outcomes, but a substantial portion improved on the performance-based tests but not on the BASFI questionnaire.

In the light of these results, the question arises: Is the chosen cut-off point of 20% to define improvement correct? This can be divided in two sub-questions:

1. Does it refer to a (minimal) clinically important difference (MCID)?
2. Does it exceed the measurement error?

As to the first question, the MCID for the performance-based tests has yet to be determined. The MCID is often calculated using a ‘transition’ measure or anchor-based method. In this method, patients are surveyed about their perception of change that has occurred as a result of treatment, indicated on an external criterion, or anchor (e.g., a 7-point Likert scale). Patients thus rate their perceived change. However, this method is also patient-reported and can be biased by estimates of the prior state (response shift or implicit theory of change) [18]. We therefore did not consider this as an appropriate anchor for an (objective), performance-based outcome measure.

A definite reply to the second question is more challenging. However, it is unlikely that all observed changes were attributable to measurement error: the reliability of the measurements was high and the SEM was relatively small, ranging from 3.6% - 8.7% of the observed range.

Apart from being reliable, valid, and responsive, an outcome measurement should also be feasible in daily clinical practice, and redundant testing should be prevented or eliminated. In **Chapter 5**, therefore, a selection of tests was made. All tests were well tolerated by the patients, who had varying limitations in physical functioning. The performance tests were designed in such a way that patients with severe disease could perform them, but the tests were also strenuous enough to differentiate highly functional patients. Based on ICC, SRM (standardized response mean), and the percentage of improvers after 3 months of TNFi therapy, the tests for ‘bending’, ‘putting on socks’, and ‘getting up from the floor’ were selected. Patients were defined as improvers on the AS Performance-based Improvement criterion (ASPI) if they had a $\geq 20\%$ improvement on one or more of the three selected tests and no deterioration of $\geq 20\%$ on the remaining test(s). A point in favour of this definition of improvement is that consistent improvement (or absence of deterioration) has to be present in three separate tests simultaneously. Because multiple tests are taken into account, it gives greater confidence in the idea that the change is real [17].

Lastly, the percentage of improvers on the full set of tests were compared to the percentage identified with the ASPI. The results showed that the ASPI criteria produced similar information in comparison with the full set of tests. The percentage improvers on the ASPI was 67% (compared to 63% for the full set) and the percentage of improvers in ASAS20 non-responders was 18% (compared to 16% for the full set). The selection of the three tests thus generated equivalent information to that of the full set, were easy to administer, and seemed to be feasible in daily clinical practice.

Disagreement between self-reported and performance-based assessments in AS

In this thesis we described limited agreement between the performance-based tests and patient-reported questionnaires of physical functioning. This observation is frequently described in literature and it is suggested that both assessment methods measure different aspects of the same construct. As outlined above, self-reported measures refer to what a person perceives he/she can do, which acknowledges the perspective and prior experiences of the patient [19]. Performance-based outcomes measure what a patient actually can do and are free of bias that can be introduced by persons making implicit judgments. The debate on clarifying the differences between both outcome measures is still ongoing. Numerous factors have been claimed to contribute to the moderate associations between the methods. Below some of these factors are outlined.

First, to complete a questionnaire, patients are required to understand the question and recall the relevant behaviour [18]. Cognitive impairments, culture, language, and education may influence the response [10-12]. Moreover, it has been suggested that asking a patient to recall how he/she felt last week is mostly influenced by the most severe episode of symptoms experienced just prior to filling in the questionnaire, especially in disorders that may fluctuate over time, like arthritis [18]. A study by Berthelot and colleagues indeed showed a high week-to-week variability in BASFI and BASDAI scores in spondyloarthritis patients [20]. In the studies described in this thesis, the BASDAI and BASFI were only assessed before and after 3 months of TNFi therapy and it is therefore not possible to give information on the week-to-week variation and possible influence thereof on physical functioning.

Factors related to the mental state of a patient (e.g., depression, passive coping, helplessness, and fear of movement) can also significantly influence the variability in self-reported functioning [7, 21] or influence a patient's score on a performance test. With chronic low-back pain, patients who reported a high degree of fear of movement or (re)injury showed more fear and escape/avoidance when exposed to a simple movement than patients with a low degree of fear of movement or (re)injury [22]. Encouragement by the assessor can positively affect the performance of a patient undergoing a performance test [23]. In order to prevent bias by encouragement, careful standardisation of the execution of the test is needed. In the performance tests for AS there was no encouragement. However, neither type of outcome measure reflects adaptations made in everyday living, can distinguish unmotivated from incapable patients, or accounts for personal preferences [24].

Another factor of influence on the association between self-report and performance-based assessments might relate to the specific phrasing and anchors of a questionnaire. In the BASFI questionnaire, patients are asked to indicate their level of ability (from 'easy' to 'impossible') to perform an activity. We showed that patients seem to include pain and exertion on doing a task within their assessment. They might take this into account subconsciously, but also consciously as part of the 'ease' of doing a task. It would be interesting to examine how patients interpreted the wording of the anchors of the BASFI questionnaire.

Moreover, it has been suggested that time alone may be inadequate to represent the concept of physical function [14, 15]. The performance of a task, with elapsed time serving as a surrogate for the amount of difficulty, might not cover the breadth of domains perceived

as relevant by the patient. It has been suggested that a composite score based on time, pain, and exertion might be a better approach [15]. In a clinical setting, the use of all scores (i.e., time, pain or exertion) might be the preferred approach, because a change in any of the parameters could be relevant in the evaluation of therapies. However, in a research setting it is important to only measure the intended concept (i.e., actual physical function) free from any bias caused by other parameters (i.e., pain and exertion). Furthermore, pain and exertion experienced during performance are, just like the BASFI questionnaire, self-reported measures. A composite score partly based on self-report is thus deemed to suffer from the same disadvantages: the risk of under- or overestimation. Using a composite score would, therefore, undermine the principal reason why the performance-based measures for AS based on time were developed.

The discordance between the outcome measures can also be influenced by AS-specific factors that are more strongly related to (improvement on) the performance tests on the one hand, and on the other hand, the BASFI questionnaire and ASAS20 response. Based on our findings, we hypothesise that performance-based physical functioning is more strongly based on structural damage, and self-reported physical functioning on pain and fatigue. We did find a moderate association between spinal mobility (BASMI) and the performance-based tests. Although spinal mobility and structural damage cannot be used interchangeably [25], there is some interrelatedness [26]. It would be particularly interesting to look at hip involvement, because hip involvement is known to be of great influence on physical functioning and too often be affected in AS patients [27]. This view is strengthened by hip mobility seeming to be of vital importance in all three selected ASPI tests.

Pain and fatigue could be important factors influencing self-reported physical functioning in AS patients. Fatigue is one of the predominant symptoms of AS patients. AS patients are more fatigued than healthy people [28] and fatigue is mostly associated with pain, but also with disease activity and (self-reported) physical functioning [29]. We also found the BASFI questionnaire to be more strongly related to exertion and pain associated with executing the performance tests.

The model for health outcomes in AS by Machado et al. [30] describes that (self-reported) physical function is determined by both disease activity (i.e., pain, fatigue, etc.) and spinal mobility. Spinal mobility impairment, in turn, is determined by irreversible spinal damage and by reversible spinal inflammation, with spinal mobility impairment being more influenced by spinal inflammation in early disease, and by structural damage in later disease [30-31]. The information on structural damage, radiographic progression, hip involvement, and fatigue was not available in our studies and this relations needs to be evaluated in future.

It would also be worth knowing if there is a stronger association between improvers on the performance tests and the activity level of patients. It has been shown that TNFi therapy has a positive effect on the activity level of patients. After TNFi therapy, patients train more often, and at a higher intensity, which could lead to a greater improvement in performance-based physical functioning [32]. The role of physical activity in relation to performance-based or self-reported physical functioning needs to be further clarified, however.

Finally, it would be interesting if improvement on performance-based (and self-reported) physical functioning could be predicted by certain patient characteristics such as age, gender, disease duration, BASFI score, etcetera. However, in our relatively small study population, we could not ascribe the improvers after TNFi therapy with any specific characteristics. Age, disease duration, and gender were not found to be associated with improvement in performance-based physical functioning. Although, the non-improvers in performance-based physical function did show slightly lower baseline values for performance, only this difference did not reach statistical significance. However, it could be that patients without performance-based improvement experienced a milder impact on physical function and could therefore have had less chance for improvement. Maybe in future research, using larger populations, positive predictors of physical improvement could be described.

Prediction of physical functioning in Ankylosing Spondylitis patients

Currently, only limited information is available on the long-term course, outcome, and predictors of limitations in physical functioning in AS patients receiving TNFi therapy in daily clinical practice. More insight into this could aid the selection of patients or give direction to additional treatment options.

In **Chapter 6** the 3-year course, outcome, and predictors of self-reported physical functioning (BASFI) were studied in an observational cohort of AS patients eligible for TNFi treatment. Results showed that physical functioning gradually improved up to 6 months of TNFi treatment and stabilised thereafter.

According to the ASAS/EULAR recommendations, the time of evaluation of the efficacy of TNFi treatment should be between 6 and 12 weeks [33], but, in view of the results of the present thesis, a better period to evaluate efficacy would be after 6 months. In clinical practice this evaluation is often taking place after 6 months, and therefore an adaptation of the guidelines may be necessary in future.

Furthermore, we found that lower baseline BASFI- and BASMI-scores, absence of comorbidity, younger age, and lower body mass index (BMI) were predictors of a more favourable outcome and course of physical functioning after 3 years of TNFi therapy. This is in line with previous studies. Also Durcan et al. showed that overweight AS patients have more limitations in physical functioning [34], but the negative influence of higher BMI on physical functioning has not previously been described in long-term observational TNFi cohorts. Moreover, the clinical response to TNFi could be negatively influenced by high BMI due to a higher TNF alpha level in the blood produced by fatty tissues. Higher doses of TNFi might thus be necessary in overweight AS patients [35].

In clinical practice it is of paramount importance to prevent (further) functional decline. Up-to-date knowledge of risk factors for functional decline is therefore needed. Despite the fact that it remains difficult to predict the course and outcome of the disease, the data described in Chapter 6 of this thesis can be helpful to stimulate patients to adopt an active lifestyle and/or weight reduction.

Suggestions for future research

The performance-based tests described in this thesis seem to be a promising outcome measure in the assessment of physical functioning in AS patients. However, reliability, validity, and responsiveness were assessed in a single, relatively small population and in an observational study. This influences, for instance, the measurement error and limits the generalizability. In addition, we do not know whether the observed change remained stable over time and whether patients not receiving TNFi therapy might also have changed. To make these tests useful in future, we need to replicate these results in a larger, placebo-controlled trial with a longer follow-up duration.

Secondly, we need to establish whether these tests are sensitive enough to measure changes after non-pharmacological interventions. The two main components of AS treatment are medicine and (supervised) exercise therapy. We now have an alternative outcome to assess physical functioning after TNFi therapy, but not yet after interventions where the treatment of limitations in physical functioning is the specific and main goal. It would therefore be worthwhile to assess the reliability and responsiveness of these performance-based tests in, for example, a physiotherapy setting.

Thirdly, the selection of tests (i.e., ASPI) should be assessed in clinical practice and in a broader spectrum of patients (e.g., in patients with nr-axial SpA). This information will improve the quality and feasibility of the performance-based tests, which in turn will enhance the evaluation of treatment effects of various interventions on physical functioning in AS patients.

Lastly, it would be interesting to examine whether performance-based tests are sensitive enough to identify early declines in a high-functioning AS population. It has been suggested that performance-based measures could identify information that might be predictive of early problems in physical functioning (i.e., what has been called 'preclinical disability'), before they are identified by self-report [14, 36]. Patients with AS should be diagnosed and effectively treated as early as possible and preferably before any limitations in physical functioning or structural damage have occurred [37]. Identifying early decline is thus important, and interventions to prevent functional decline should therefore start as early as possible. Patients with a longer disease duration may adapt to a certain amount of decline in function [13] or deny problems [38]. Additionally, the baseline level of the BASFI questionnaire is an important predictor of an unfavourable course and outcome of physical functioning after 3 years of TNFi therapy. Intervening at an early stage and preserving the level of physical functioning or preventing co-morbidities may even turn out to be more effective and less costly in the long run, but this remains to be investigated.

Recently, the ASAS has endorsed this project, and we hope that future collaboration with other research centres will enable to achieve the aforementioned research objectives.

Implications for clinical practice

The work described in this thesis supports the use of performance-based measures as an outcome measure for the assessment of physical functioning in patients with AS. This is a step forward towards a feasible and alternative assessment method. We want to propagate these tests as a different approach to the assessment of physical functioning in AS, not as a better assessment method. The methods are complementary, not opposing. The choice of which outcome parameter to use depends on time, money, context (research or clinical practice), intended construct (the performance tests provide information about what a patient *can* do and the BASFI questionnaire provides information about what a patient *perceives* he/she can do), and purpose of measurement (discriminative, predictive, or evaluative). Generally, for a comprehensive assessment a blend of both types of measurement is desirable, because each measure provides unique and useful information complementary to the other.

Based on the results in this thesis, the use of performance-based tests (i.e., ASPI) is recommended for both research and clinical practice for the evaluation of treatment. In order to provide a comprehensive assessment of physical functioning in AS patients, the use of the ASPI alongside the BASFI questionnaire is recommended. In the evaluation of response to TNFi therapy the performance tests can also be compared with the ASDAS (AS disease activity index). The ASDAS combines objective information (CRP or ESR) with four self-reported questions regarding AS-specific symptoms to define disease activity. By combining the ASDAS with performance-based physical functioning (i.e., the ASPI), almost all relevant domains targeted by TNFi therapy will be fully covered and evaluated efficiently.

Rheumatologists, being the experts who coordinate the treatment of AS patients, who make the decision on starting pharmacological and/or non-pharmacological treatment, and who monitor the patient (disease), should use the ASPI. Their decisions are based on and tailored according to manifestations, symptoms, prognostic indicators, and clinical status. It is therefore important to have adequate information on all relevant domains, including physical function. To diagnose and evaluate this domain swiftly and comprehensively, the performance-based tests (ASPI) can be of added value.

Physical therapists should also use the ASPI, because the physical functioning domain is their main scope and field of application. During therapy, physical therapists constantly evaluate what a person can or cannot do. In the ASAS core-set for physical therapy, however, only outcome measures assessing what a patient perceives he/she can do are included. Performance-based test of functioning can therefore be a valuable asset. In the physiotherapy setting it might be good to also use a simple test for the assessment of cardiorespiratory fitness in addition to the ASPI. AS patients have lower cardiorespiratory fitness levels, lower lung function and a decrease in exercise tolerance [39-41]. These disorders are commonly part of limitations in physical functioning, and consequently treatment will be aimed at this. Assessment of cardiorespiratory fitness in addition to the ASPI is therefore proposed for physiotherapy setting.

Finally, to increase the feasibility of the performance-based test for clinical practice, three tests were selected (i.e., ASPI). The execution of the tests appears possible within 15 minutes and can be easily executed under supervision by, for instance, a research nurse or physical therapist. The feasibility in national and international context should always be a focus of attention to warrant good dissemination and implementation of the ASPI.

Conclusions and recommendations

- Performance-based tests of physical functioning in AS appear to be reliable, valid, and responsive.
- The BASFI questionnaire and performance-based tests measure different, aspects of the physical functioning domain. This suggests both methods to be complementary rather than overlapping. Both add to the understanding of physical functioning in AS patients.
- After 3 months of TNFi treatment, 63% of the patients showed a significant improvement of $\geq 20\%$ in performance-based physical function (i.e., time).
- A selection of three tests is sufficient to measure limitations in performance-based physical functioning in patients receiving TNFi therapy.
- The level of physical functioning at the start of TNFi therapy is an important predictor of a favourable outcome after 3 years of treatment.
- In research and clinical practice, performance-based physical functioning should be assessed alongside self-reports of function and disease activity.
- More research is necessary in order to improve the quality, responsiveness, and feasibility of the performance-based tests in different treatment settings and after longer periods of time.

References

1. 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(12): 2281-5.
2. Piva SR, Fitzgerald GK, Irrgang JJ, Bouzubar F, Starz TW. Get up and go test in patients with knee osteoarthritis. *Arch Phys Med Rehabil* 2004, 85(2): 284-9.
3. Sciruba F, Criner GJ, Lee SM, Mohsenifar Z, Shade D, Slivka W, et al. National Emphysema Treatment Trial Research Group. Six-minute walk distance in chronic obstructive pulmonary disease: reproducibility and effect of walking course layout and length. *Am J Respir Crit Care Med* 2003, 167(11): 1522-7.
4. Singh SJ, Morgan MD, Scott S, Walters D, Hardman AE. Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax* 1992, 47(12): 1019-24.
5. Morales FJ, Montemayor T, Martinez A. Shuttle versus six-minute walk test in the prediction of outcome in chronic heart failure. *Int J Cardiol* 2000, 76(2-3): 101-5.
6. Kennedy DM, Stratford PW, Wessel J, Gollish JD, Penney D. Assessing stability and change of four performance measures: a longitudinal study evaluating outcome following total hip and knee arthroplasty. *BMC Musculoskeletal Disord* 2005, 6: 3.
7. Wittink H, Rogers W, Sukiennik A, Carr DB. Physical functioning: self-report and performance measures are related but distinct. *Spine* 2003, 28(20): 2407-13.
8. Hoeymans N, Wouters ER, Feskens EJ, van den Bos GA, Kromhout D. Reproducibility of performance-based and self-reported measures of functional status. *J Gerontol A Biol Sci Med Sci* 1997, 52(6): M363-8.
9. 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(7): 737-45.
10. Elam JT, Graney MJ, Beaver T, el Derwi D, Applegate WB, Miller ST. Comparison of subjective ratings of function with observed functional ability of frail older persons. *Am J Public Health* 1991, 81(9): 1127-30.
11. Sager MA, Dunham NC, Schwantes A, Mecum L, Halverson K, Harlowe D. Measurement of activities of daily living in hospitalized elderly: a comparison of self-report and performance-based methods. *J Am Geriatr Soc* 1992, 40(5): 457-62.
12. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol* 1994, 49(2): M85-94.
13. 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(10): 1103-10.
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(7): 724-31.
15. Stratford PW, Kennedy D, Pagura SM, Gollish JD. The relationship between self-report and

- performance-related measures: questioning the content validity of timed tests. *Arthritis Rheum* 2003, 49(4): 535-40.
16. 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 back pain, hand osteoarthritis, and hip and knee osteoarthritis: Results from a prospective multinational study. *Arthritis Care Res (Hoboken)* 2012, 64(11): 1699-707.
 17. 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(8): 1876-86.
 18. Streiner DL, Norman GR, Health measurement scales, Hamilton, Ontario, Canada: Oxford university press, 2003, p283.
 19. Dagfinrud H, Vollestad NK, Loge JH, Kvien TK, Mengshoel AM. Fatigue in patients with ankylosing spondylitis: A comparison with the general population and associations with clinical and self-reported measures. *Arthritis Rheum* 2005, 53(1): 5-11.
 20. Berthelot JM, Tortellier L, Lavy-Bregeon D, Le Goff B, Maugars Y. High intraindividual week-to-week variability in BASDAI and BASFI values: are several evaluations needed before starting or stopping TNFalpha antagonist therapy for spondyloarthropathies? *Joint Bone Spine* 2008, 75(2): 167-71.
 21. Brionez TF, Assassi S, Reveille JD, Leach TJ, Diekman L, Ward MM, et al. Psychological correlates of self-reported functional limitation in patients with ankylosing spondylitis. *Arthritis Res Ther* 2009, 11: R182.
 22. Vlaeyen JW, Kole-Snijders AM, Boeren RG, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain* 1995, 62(3): 363-72.
 23. Guyatt GH, Pugsley SO, Sullivan MJ, Thompson PJ, Berman L, Jones NL, et al. Effect of encouragement on walking test performance. *Thorax* 1984, 39(11): 818-22.
 24. Myers AM, Holliday PJ, Harvey KA, Hutchinson KS. Functional performance measures: are they superior to self-assessments? *J Gerontol* 1993, 48: M196-M206.
 25. Wanders A, Landewé R, Dougados M, Mielants H, van der Linden S, van der Heijde D. Association between radiographic damage of the spine and spinal mobility for individual patients with ankylosing spondylitis: can assessment of spinal mobility be a proxy for radiographic evaluation? *Ann Rheum Dis* 2005, 64(7): 988-94.
 26. Machado P, Landewé R, Braun J, Hermann KG, Baker D, van der Heijde D. Both structural damage and inflammation of the spine contribute to impairment of spinal mobility in patients with ankylosing spondylitis. *Ann Rheum Dis* 2010, 69(8): 1465-70.
 27. Vander Cruyssen B, Vastesaeger N, Collantes-Estévez E. Hip disease in ankylosing spondylitis. *Curr Opin Rheumatol* 2013, 25(4): 448-54.
 28. Alkan BM, Fidan F, Erten Ş, Aksekili H, Alemdar A, Eroğlu E, et al. Fatigue and correlation with disease-specific variables, spinal mobility measures, and health-related quality of life in ankylosing spondylitis. *Mod Rheumatol* 2013, 23(6): 1101-7.
 29. Brophy S, Davies H, Dennis MS, Cooksey R, Husain MJ, Irvine E, et al. Fatigue in ankylosing spondylitis: treatment should focus on pain management. *Semin Arthritis Rheum* 2013, 42(4): 361-7.
 30. Machado P, Landewé R, Braun J, Hermann KG, Baraliakos X, Baker D, et al. A stratified model for health outcomes in ankylosing spondylitis. *Ann Rheum Dis* 2011, 70(10): 1758-64.
 31. 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 (6): 863-7.
32. Dubey SG, Leeder J, Gaffney K. Physical therapy in anti-TNF treated patients with ankylosing spondylitis. *Rheumatology (Oxford)* 2008, 47(7): 1100-1.
 33. Zochling J, van der Heijde D, Burgos-Vargas R, Collantes E, Davis JC Jr, Dijkmans B, et al. 'Assessment in AS' international working group, European League Against Rheumatism. ASAS/EULAR recommendations for the management of ankylosing spondylitis. *Ann Rheum Dis* 2006, 65(4): 442-52.
 34. Durcan L, Wilson F, Conway R, Cunnane G, O'Shea FD. Increased body mass index in ankylosing spondylitis is associated with greater burden of symptoms and poor perceptions of the benefits of exercise. *J Rheumatol* 2012, 39: 2310-4
 35. Ottaviani S, Allanore Y, Tubach F, Forien M, Gardette A, Pasquet B, et al. Body mass index influences the response to infliximab in ankylosing spondylitis. *Arthritis Res Ther* 2012, 14(3): R115.
 36. Fried LP, Bandeen-Roche K, Chaves PH, Johnson BA. Preclinical mobility disability predicts incident mobility disability in older women. *J Gerontol A Biol Sci Med Sci* 2000, 55(1): M43-52.
 37. Sieper J, Appel H, Braun J, Rudwaleit M. Critical appraisal of assessment of structural damage in ankylosing spondylitis: implications for treatment outcomes. *Arthritis Rheum* 2008, 58(3): 649-56.
 38. Hidding, A, van der linden S, Boers M, Gielen X, Kester A, [Vlaeyen J](#). Fake good test-taking attitude in ankylosing-spondylitis. *Arthritis and Rheum* 1992, 35(9): S244-S244.
 39. Halvorsen S, Vøllestad NK, Fongen C, Provan SA, Semb AG, Hagen KB, et al. Physical fitness in patients with ankylosing spondylitis: comparison with population controls. *PhysTher* 2012, 92(2): 298-309.
 40. Brambila-Tapia AJ, Rocha-Muñoz AD, Gonzalez-Lopez L, Vázquez-Del-Mercado M, Salazar-Páramo M, Dávalos-Rodríguez IP, et al. Pulmonary function in ankylosing spondylitis: association with clinical variables. *Rheumatol Int* 2013, 33(9): 2351-8.
 41. Halvorsen S, Vøllestad NK, Provan SA, Semb AG, van der Heijde D, Hagen KB, et al. Cardiorespiratory fitness and cardiovascular risk in patients with ankylosing spondylitis: a cross-sectional comparative study. *Arthritis Care Res (Hoboken)* 2013, 65(6):969-76.