External validation of the alternative Ankylosing Spondylitis Disease Activity Score in three randomized clinical trials of ixekizumab

Ortolan, A.; Ramiro, S.; Ramonda, R.; Heijde, D. van der

Rheumatology, 2022, **00**, 1–5
https://doi.org/10.1093/rheumatology/keac618
Advance access publication 25 October 2022
Concise report

British Society for Rheumatology RHEUMATOLOGY OXFORD

Clinical science

External validation of the alternative Ankylosing Spondylitis Disease Activity Score in three randomized clinical trials of ixekizumab

Augusta Ortolan [1,2], Sofia Ramiro [1,3,*], Roberta Ramonda [2], Désirée van der Heijde [1]

[1]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[2]Rheumatology Unit, Department of Medicine, University of Padova, DIMED, Padova, Italy
[3]Department of Rheumatology, Zuyderland Medical Center, Heerlen, The Netherlands

*Correspondence to: Sofia Ramiro, Department of Rheumatology, Leiden University Medical Center, Albinusdreef 2, 2333 GA Leiden, The Netherlands, P.O. Box 9600, 2300RC Leiden, The Netherlands. E-mail: sofiaramiro@gmail.com

Abstract

Objectives: The alternative ASDAS (altASDAS) is an index that can be used when patient global assessment is unavailable. Our aim was to test the truth and discrimination aspects according to OMERACT filter 2.0 of the altASDAS in an external cohort.

Methods: Cohorts from the COAST trials of ixekizumab (COAST-V, -W, -X; 16-week primary endpoint) enrolling radiographic/non-radiographic axial SpA patients were pooled. The ASDAS [original formula with patient global assessment (PGA)] and altASDAS were calculated. Truth was assessed by agreement with the continuous ASDAS [intraclass correlation coefficients (ICCs)] and ASDAS disease activity (DA) states (weighted κ), Bland–Altman plots [mean difference (MD) and 95% limits of agreement (LoA)] and Pearson's correlations between altASDAS/ASDAS and other constructs. Discrimination was tested by the ability of altASDAS to distinguish high/low DA according to nocturnal pain >6/10 as an external anchor and agreement (κ) with ASDAS in major improvement (MI) and clinically important improvement (CII).

Results: A total of 958 patients were included. For truth, agreement with ASDAS was very good (ICC = 0.99, κ = 0.91), MD with ASDAS was 0.03 (95% LoA −0.31–0.24) and correlation coefficients of altASDAS with related constructs were within a prespecified 0.3-wide band around those between ASDAS and the same construct. For discrimination, the altASDAS discriminated between DA states and agreed with ASDAS response (κ MI = 0.91, CII = 0.93).

Conclusions: The altASDAS was truthful and discriminative in an external cohort and as such has been fully validated to be used in cases when PGA is unavailable.

Keywords: axial spondyloarthritis, disease activity, Ankylosing Spondylitis Disease Activity Score (ASDAS), BASDAI, validation

Rheumatology key messages

- When patient global assessment is unavailable, this can be replaced by the BASDAI total to calculate the alternative ASDAS (altASDAS).
- The truth and discrimination of the altASDAS have been confirmed in an external cohort.
- The altASDAS is truthful, discriminative and feasible and can be used in datasets when patient global assessment is unavailable.

Introduction

In axial spondyloarthritis (axSpA), disease activity is assessed by the Ankylosing Spondylitis Disease Activity Score (ASDAS), a weighted instrument incorporating patient-reported outcomes (PROs) and an inflammation marker (preferably CRP) [1, 2]. This index has been developed to overcome the limitations of the fully patient-reported, and unweighted, BASDAI, which consists of six questions. Three of these questions (Q2, back pain; Q3, peripheral joint pain/swelling; and Q6, duration of morning stiffness) are also included in the ASDAS, in addition to the Patient Global Assessment of disease activity (PGA) [2]. The ASDAS has better validity, discriminatory capacity and sensitivity to change than the BASDAI [3]. Since the ASDAS was introduced later than the BASDAI, it was often the case that in axSpA cohorts or registries, PGA was not collected, meaning that the ASDAS could not be calculated. In order to overcome this problem, we previously developed and validated an 'alternative' ASDAS (altASDAS) to be used when PGA is unavailable [4]. When calculating the altASDAS, the PGA is replaced by the BASDAI total score, which results in maintaining the closest psychometric properties to the original index (hereby referred to as 'ASDAS').

Received: 28 July 2022. **Accepted:** 15 October 2022

So far, the altASDAS has only been validated in one cohort [4]. In order to fully validate it, however, it is necessary to test its performance in an independent cohort. A first attempt at this task was made by Aranda-Valera *et al.* [5, 6], who were also the developers of another alternative index, the 'BASDAS', which was created as a surrogate for the ASDAS when the separate BASDAI questions are not available. However, in their validation cohort, the agreement between the altASDAS and ASDAS was assessed, but not other important aspects such as discrimination, especially regarding the score's change over time [6]. Therefore, the aim of the present study was to test the truth and discrimination aspects of the altASDAS in an external cohort [7] according to the OMERACT filter 2.0, while feasibility from our previous work was considered [4].

Methods

Study population

Patients with both radiographic and non-radiographic axSpA (r-axSpA and nr-axSpA) participating in the COAST trial program of ixekizumab (COAST-V, COAST-W, COAST-X) were considered for this analysis [8–10]. COAST-V and COAST-W enrolled adult r-axSpA patients, fulfilling Assessment of SpondyloArthritis international Society (ASAS) classification criteria [8, 9]. Patients were TNF inhibitor (TNFi) naïve and TNFi experienced, respectively. In contrast, COAST-X enrolled adult nr-axSpA patients fulfilling ASAS classification criteria [10]. In all trials, patients had active disease, defined as a BASDAI $\geq$4 and total back pain numeric rating scale $\geq$4, both at screening and baseline. Furthermore, in all three randomized controlled trials (RCTs), two dose regimens of ixekizumab (80 mg every 2 weeks and 80 mg every 4 weeks) were compared with a placebo arm, and the primary endpoint was an ASAS 40% response at week 16 [8–10]. For the purpose of this validation study, both ixekizumab dosages were considered together as the 'treatment arm', which was compared with the placebo arm. Assessments were carried out at weeks 0 (baseline), 4, 8 and 16.

Ethics committee approval was obtained for the RCTs COAST-V (NCT02696785), COAST-W (NCT02696798) and COAST-X (NCT02757352) by each participating centre. No additional data were required for the present analysis.

Assessments

At all time points, the BASDAI total score and its individual questions on fatigue (Q1), back pain (Q2), peripheral joint pain/swelling (Q3), enthesitis (Q4), severity (Q5) and duration of morning stiffness (Q6) were considered. Their value ranged from 0 (none) to 10 (very severe). The ASDAS was calculated with the usual formula: $0.12 \times Q2 + 0.06 \times Q6 + 0.11 \times PGA + 0.07 \times Q3 + 0.58 \times \ln(CRP + 1)$. At week 16, the fulfilment of ASDAS clinically important improvement (ASDAS-CII; decrease $\geq$1.1 units) and major improvement (ASDAS-MI; decrease $\geq$2 units) was also assessed.

In addition, the altASDAS was calculated at all time points according to the formula: $0.12 \times Q2 + 0.06 \times Q6 + 0.11 \times 0.99 \times$ BASDAI total score $+ 0.07 \times Q3 + 0.58 \times \ln(CRP + 1)$ [4]. Response criteria (MI = decrease $\geq$2 units, CII = decrease $\geq$1.1 units) were also calculated using the altASDAS.

For both the ASDAS and altASDAS, CRP values <2 mg/l in the formula were substituted with CRP = 2 mg/l, as recommended [11].

In order to test truth we assessed the following constructs, which are considered to be related to disease activity: physical function, by the BASFI, from 0 (no impairment) to 10 (maximum impairment); health-related quality of life (HRQoL), by the mental and physical component summaries (MCS and PCS) of the Short Form-36 questionnaire (0–100, with higher scores representing better states) [12] and overall functioning and health, by the ASAS Health Index [ASAS HI; from 0 (no impairment) to 17(maximum impairment)] [13].

Discrimination was tested using an external anchor to define a high or low disease activity state: nocturnal pain, measured by a numerical rating scale (NRS) from 0 to 10, dichotomized with an arbitrary cut-off (>6 for high and $\leq$6 for low disease activity state).

Statistical analysis

As for development of the altASDAS and its primary validation, the ASDAS was taken as the gold standard measure. Thus, firstly, truth was assessed by agreement with the ASDAS, both as a continuous score, with intraclass correlation coefficients (ICCs) using a two-way random effect model, and as a categorical variable (ASDAS disease activity states) using a weighted κ. Second, Bland–Altman plots with 95% limits of agreement (LoA) comparing ASDAS and altASDAS were created and systematic error [mean difference (MD)] and random error (scedasticity of the plot) were assessed. Finally, a comparison of Pearson's correlation coefficients between the ASDAS and related constructs (BASFI, MCS, PCS, ASAS HI) with correlation coefficients between the altASDAS and the same related constructs was performed. It was predefined that correlation coefficients between the altASDAS/related constructs had to fall within a 0.3-wide band around the corresponding correlation coefficients between the ASDAS/related construct.

Discrimination was assessed by comparing the ability of the ASDAS and altASDAS to distinguish between high and low disease activity states according to the external anchor nocturnal pain. Thus standardized mean differences (SMDs) between low and high disease activity (the difference in the means of the two groups divided by the pooled SD of the group means) for both the ASDAS and altASDAS were calculated, with higher SMDs meaning higher discrimination. The aspect of discrimination concerning change over time was tested by agreement (κ) of the altASDAS with the ASDAS in the percentages of patients achieving ASDAS-CII and ASDAS-MI at 16 weeks in the treatment *vs* placebo arm and a comparison of the number of patients reaching ASDAS-MI and ASDAS-CII in the treatment *vs* placebo arm at week 16, according to the ASDAS and altASDAS (higher χ^2 means better discrimination).

No imputation was performed for missing data, as the percentage of missing data for the ASDAS, BASDAI and separate components was <3%. Patients with missing data were simply excluded.

Statistical analyses were performed using STATA/SE version 17 (StataCorp, College Station, TX, USA). For agreement analyses, prespecified desired levels of values $\geq$0.8 for ICC and κ were established.

Results

A total of 958 patients were included, of which 671 (70%) were males, with a mean age of 43 years (SD. 12). Among

these, 656 (68%) were patients with r-axSpA. The main baseline characteristics of the population are outlined in Supplementary Table S1, available at *Rheumatology* online.

Agreement of the altASDAS with the ASDAS was excellent both in terms of continuous score [ICC 0.99 (95% CI 0.99, 0.99)] and according to disease activity states, with a weighted κ of 0.91 (95% CI 0.91, 0.91). The Bland–Altman plot, depicted in Fig. 1, revealed an MD with the ASDAS of 0.03 points with a 95% LoA of −0.31–0.24. The plot showed no proportional bias and the scatter of differences was uniform (homoscedasticity). Regarding the correlation coefficients of the altASDAS with related constructs, these fell within the prespecified 0.3-wide band around those between the ASDAS and the same construct, and in fact, they were substantially overlapping (Table 1A).

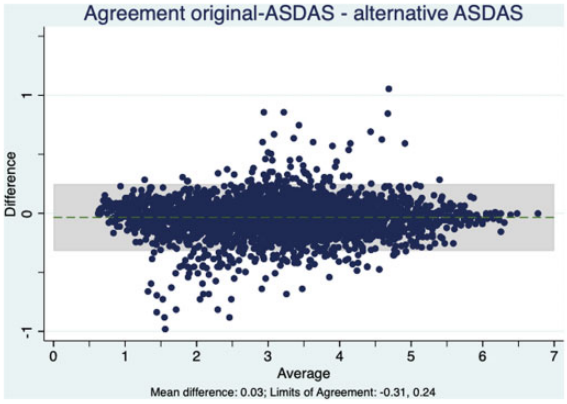

Figure 1. Bland–Altman plot with MD and 95% LoA between the ASDAS and altASDAS

Regarding discrimination, altASDAS was found to be able to discriminate between low and high disease activity states, with an SMD between the altASDAS in the groups with nocturnal pain >6 and ≤6, which was very close to the SMD of the ASDAS (Table 1B). Sensitivity to change was demonstrated by very good agreement between the altASDAS-CII and ASDAS-CII [$\kappa = 0.93$ (95% CI 0.91, 0.97)], as well as the altASDAS-MI and ASDAS-MI [$\kappa = 0.91$ (95% CI 0.88, 0.95)]. In addition, the percentages of patients reaching CII and MI were almost the same for altASDAS and ASDAS, with very similar χ^2 values, indicating an equivalent discriminatory ability (Table 1C). At week 16, the SMD in the altASDAS change and ASDAS change between the treatment and placebo arms were numerically very close, again confirming good sensitivity to change of the alternative formula (Table 1D).

Discussion

The psychometric properties (truth and discrimination) of altASDAS were tested in a different cohort from the one where the index was developed, meaning that an external validation was performed: the instrument proved to be truthful and discriminative, besides being feasible, as already discussed in our previous paper [4].

Although the ASDAS in its original form is the gold standard to measure disease activity in axSpA, it is useful to have a second option in cases where the ASDAS cannot be calculated, which often happens because the PGA is not available in existing datasets. It might be argued that in these cases, using the BASDAI is an equally valid and already widely available alternative for the ASDAS. However, this approach has serious shortcomings. In fact, the BASDAI is an entirely

Table 1. Psychometric properties of the altASDAS

(A) Truth: correlations with constructs related to disease activity

Measure	Function (BASFI)	HRQoL (MCS)	HRQoL (PCS)	Overall functioning and health (ASAS HI)
ASDAS, *r*	0.58	−0.11	−0.48	0.35
altASDAS, *r*	0.59	−0.12	−0.48	0.37

(B) Discrimination: disease activity states

Anchor to define disease activity: nocturnal pain (0–10)	NRS >6, mean ASDAS (S.D.)	NRS ≤6, mean ASDAS (S.D.)	SMD
ASDAS	3.97 (0.76)	2.60 (0.86)	−1.68
altASDAS	3.93 (0.77)	2.57 (0.87)	−1.65

(C) Discrimination: sensitivity to change [1]

Variable	Treatment, *n* (%)	Placebo, *n* (%)	χ^2
Patients	628	272	–
ASDAS-MI	150 (24)	12 (4)	48.8
altASDAS-MI	152 (24)	13 (5)	47.8
ASDAS-CII	341 (54)	62 (23)	76.2
altASDAS-CII	343 (55)	64 (23)	74.0

(D) Discrimination: sensitivity to change [2]

Variable	Treatment, mean ASDAS change (S.D.) at week 16	Placebo, mean ASDAS change (S.D.) at week 16	SMD
Patients	628	272	–
ASDAS	1.27 (1.10)	0.41 (0.97)	−0.82
altASDAS	1.26 (1.10)	0.43 (0.97)	−0.80

subjective index, it is not weighted and, most of all, only an arbitrary cut-off (BASDAI $\geq$4) is available to dichotomously define (high and low) disease activity states [14]. This means that, using the BASDAI, it would not be possible to discriminate, for example, between inactive disease and low disease activity as can be done with the ASDAS [15, 16]. Moreover, use of the $\geq$4 threshold has been criticized, as it has been demonstrated to be a poor predictor of response to TNF inhibitors, while the ASDAS is known to better predict response to treatment [17, 18]. Furthermore, for the BASDAI, a minimum clinically important improvement (MCII) has been defined, but this has not been widely used in clinical trials or settings where patients with high disease activity initiated biologic therapy [19]. Therefore, having the option to calculate an ASDAS surrogate (altASDAS) with very good psychometric properties, and to use the same cut-offs as the ASDAS for both disease activity states and improvement, is valuable and increases the comparability across studies.

Regarding its psychometric properties, the altASDAS showed excellent agreement with the ASDAS continuous score and disease activity states [4]. Also, the correlation coefficients between the altASDAS and related constructs (functional ability, HRQoL, overall health and function) were practically equal to those between the ASDAS and the same constructs, confirming that the altASDAS is a truthful measure for disease activity.

The discriminatory ability of the altASDAS was also extremely close to the ASDAS, both for status and change scores. The latter is particularly important because it differentiates the altASDAS form other alternative indices, such as the BASDAS [5]. The BASDAS, which makes use of the BASDAI total score and CRP—and not individual BASDAI questions or PGA—demonstrated very good agreement with the ASDAS (ICC = 0.96), much like the altASDAS (ICC = 0.99), albeit with a wider LoA [6]. In addition, agreement in disease activity states was good for both the BASDAS ($\kappa = 0.88$) and the altASDAS ($\kappa = 0.91$). However, the BASDAS sensitivity to change has not been tested so far. Thus altASDAS is the only index for which the cut-offs for CII and MI have been compared with the ASDAS, and have demonstrated very good agreement. This will allow researchers to use the altASDAS in longitudinal research as well. Furthermore, psychometric performances were in general slightly better for the altASDAS than the BASDAS because, while the altASDAS was selected as the best-performing index among a variety of substitutes, the BASDAS was formulated deciding a priori to use the BASDAI total score and CRP, as the aim was different (i.e. to develop an index to be used when individual BASDAI questions are not available) [20].

The main limitation of the present study is the use of RCT populations including only patients with active disease (BASDAI $\geq$4), which may have caused underrepresentation of patients with low or inactive disease at baseline. On the other hand, considering a population where a new therapy was initiated in a strictly controlled environment was necessary in order to optimally test sensitivity to change. The strengths of the work are the heterogeneity of the population, including both r-axSpA and nr-axSpA, the very low percentage of missing data throughout evaluations, as well as following the methodology endorsed by OMERACT and comprehensively assessing all psychometric properties relevant to ensure one instrument has them before its widespread use.

In conclusion, the altASDAS was shown to be truthful and discriminative in an external cohort and as such has been fully validated and can be used in cases when the PGA is unavailable in existing datasets.

Supplementary data

Supplementary data are available at *Rheumatology* online.

Data availability statement

Request for use should be submitted to the original owner of the databases (Eli Lilly).

Funding

A.O. received a grant from Eli Lilly for this work through the 2020 New Horizon Fellowship.

Disclosure statement: A.O. received a grant from Eli Lilly through the 2020 New Horizon Fellowship. S.R. received research grants from AbbVie, Galapagos, MSD, Novartis, Pfizer and UCB and honoraria or consultancy/speaker fees from AbbVie, Eli Lilly, MSD, Novartis, UCB and Sanofi. R.R. received honoraria or consultancy/speaker fees from AbbVie, Amgen, Bristol-Meyers Squibb, Eli Lilly, Galapagos, Janssen, Novartis, Pfizer, Roche and UCB Pharma. D.v.d.H. received honoraria or consultancy/speaker fees from AbbVie, Amgen, Astellas, AstraZeneca, Bayer, Bristol-Meyers Squibb, Boehringer Ingelheim, Celgene, Cyxone, Daiichi, Eisai, Eli Lilly, Galapagos, Gilead, GlaxoSmithKline, Janssen, Merck, Novartis, Pfizer, Regeneron, Roche, Sanofi, Takeda and UCB Pharma and is Director of Imaging Rheumatology bv.

Acknowledgements

The authors wish to thank Eli Lilly for providing data of the trials COAST-V, (NCT02696785), COAST-W (NCT02696798) and COAST-X (NCT02757352) and for supporting A.O. through a grant for the 2020 New Horizon Fellowship. All patients provided written informed consent prior to any protocol-specific procedures.

References

1. van der Heijde D, Lie E, Kvien TK *et al.* ASDAS, a highly discriminatory ASAS-endorsed disease activity score in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68:1811–8.
2. Lukas C, Landewé R, Sieper J *et al.* Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68:18–24.
3. Garrett S, Jenkinson T, Kennedy LG *et al.* A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol 1994;21:2286–91.
4. Ortolan A, Ramiro S, van Gaalen F *et al.* Development and validation of an alternative ankylosing spondylitis disease activity score when patient global assessment is unavailable. Rheumatology (Oxford) 2021;60:638–48.
5. Aranda-Valera IC, Garrido-Castro JL, Ladehesa-Pineda L *et al.* How to calculate the ASDAS based on C-reactive protein without individual questions from the BASDAI: the BASDAI-based ASDAS formula. Rheumatology (Oxford) 2020;59:1545–9.
6. Aranda-Valera IC, Garrido-Castro JL, Ladehesa L *et al.* Comment on: Development and validation of an alternative ankylosing spondylitis disease activity score when patient global assessment is unavailable. Rheumatology (Oxford) 2021;60:e69–70.

7. Boers M, Kirwan JR, Wells G *et al.* Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol 2014;67:745–53.
8. van der Heijde D, Cheng-Chung Wei J, Dougados M *et al.* Ixekizumab, an interleukin-17A antagonist in the treatment of ankylosing spondylitis or radiographic axial spondyloarthritis in patients previously untreated with biological disease-modifying anti-rheumatic drugs (COAST-V): 16 week results of a phase 3 randomised, double-blind, active-controlled and placebo-controlled trial. Lancet 2018;392:2441–51.
9. Deodhar A, Poddubnyy D, Pacheco-Tena C *et al.* Efficacy and safety of ixekizumab in the treatment of radiographic axial spondyloarthritis: sixteen-week results from a phase III randomized, double-blind, placebo-controlled trial in patients with prior inadequate response to or intolerance of tumor necrosis factor inhibitors. Arthritis Rheumatol 2019;71:599–611.
10. Deodhar A, van der Heijde D, Gensler LS *et al.* Ixekizumab for patients with non-radiographic axial spondyloarthritis (COAST-X): a randomised, placebo-controlled trial. Lancet 2020;395:53–64.
11. Machado P, Navarro-Compán V, Landewé R *et al.* Calculating the ankylosing spondylitis disease activity score if the conventional c-reactive protein level is below the limit of detection or if high-sensitivity C-reactive protein is used: an analysis in the DESIR cohort. Arthritis Rheumatol 2015;67:408–13.
12. McHorney CA, Ware JE, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993;31:247–63.
13. Kiltz U, van der Heijde D, Boonen A *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. Machado P, van der Heijde D. How to measure disease activity in axial spondyloarthritis? Curr Opin Rheumatol 2011;23:339–45.
15. Machado P, Landewé R, Lie E *et al.* Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores. Ann Rheum Dis 2011;70:47–53.
16. Machado PM, Landewé R, Heijde DV, Assessment of SpondyloArthritis international Society (ASAS). Ankylosing Spondylitis Disease Activity Score (ASDAS): 2018 update of the nomenclature for disease activity states. Ann Rheum Dis 2018;77:1539–40.
17. Braun J, Baraliakos X, Kiltz U *et al.* Disease activity cutoff values in initiating tumor necrosis factor inhibitor therapy in ankylosing spondylitis: a German GO-NICE study subanalysis. J Rheumatol 2020;47:35–41.
18. Nam B, Koo BS, Lee T-H *et al.* Low BASDAI score alone is not a good predictor of anti-tumor necrosis factor treatment efficacy in ankylosing spondylitis: a retrospective cohort study. BMC Musculoskelet Disord 2021;22:140.
19. Kviatkovsky MJ, Ramiro S, Landewé R *et al.* The minimum clinically important improvement and patient-acceptable symptom state in the BASDAI and BASFI for patients with ankylosing spondylitis. J Rheumatol 2016;43:1680–6.
20. Ortolan A, Ramiro S, Bastiaenen C, van der Heijde D. Comment on: Development and validation of an alternative ankylosing spondylitis disease activity score when patient global assessment is unavailable: reply. Rheumatology (Oxford) 2021;60:e71–2.