# Chapter 2

## Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores

Machado P, Landewé R, Lie E, Kvien TK, Braun J, Baker D, van der Heijde D

ANN RHEUM DIS. 2011 JAN;70(1):47-53

# ABSTRACT


**Background**

The Ankylosing Spondylitis Disease Activity Score (ASDAS) is a new composite index to assess disease activity in ankylosing spondylitis (AS). It fulfils important aspects of truth, feasibility and discrimination. Criteria for disease activity states and improvement scores are important for use in clinical practice, observational studies and clinical trials and so far have not been developed for the ASDAS.

**Objectives**

To determine clinically relevant cut-off values for disease activity states and improvement scores using the ASDAS.

**Methods**

For the selection of cut-offs data from the Norwegian disease modifying antirheumatic drug (NOR-DMARD) registry, a cohort of patients with AS starting conventional or biological DMARDs, were used. Receiver operating characteristic analysis against several external criteria was performed and several approaches to determine the optimal cut-offs used. The final choice was made on clinical and statistical grounds, after debate and voting by Assessment of SpondyloArthritis international Society members. Cross-validation was performed in NOR-DMARD and in Ankylosing Spondylitis Study for the Evaluation of Recombinant Infliximab Therapy, a database of patients with AS participating in a randomised placebo-controlled trial with a tumour necrosis factor blocker.

**Results**

Four disease activity states were chosen by consensus: inactive disease, moderate, high and very high disease activity. The three cut-offs selected to separate these states were: 1.3, 2.1 and 3.5 units. Selected cut-offs for improvement were: change ≥1.1 units for clinically important improvement and change ≥2.0 units for major improvement. Results of the cross-validation strongly supported the cut-offs.

**Conclusion**

Cut-off values for disease activity states and improvement using the ASDAS have been developed. They proved to have external validity and a good performance compared to existing criteria.

# INTRODUCTION



Ankylosing spondylitis (AS) is a chronic inflammatory rheumatic disease that affects the axial skeleton. It is characterised by inflammatory back pain, bony fusion of the spine, decreased mobility, functional impairment and decreased quality of life. Other clinical features of AS include asymmetric peripheral oligoarthritis, enthesitis, fatigue and specific organ involvement such as anterior uveitis, psoriasis and chronic inflammatory bowel disease.[1]

The concept of disease activity, a reflection of the underlying inflammation, encompasses a wide range of domains and measures.[2] Since currently used single component measures or indices have limitations because they measure only one aspect of the disease, are fully patient or doctor oriented, or lack face and/or construct validity, the Assessment of SpondyloArthritis international Society (ASAS) has developed a new disease activity score for use in AS: the 'Ankylosing Spondylitis Disease Activity Score' (ASDAS).[3]

Designed in analogy of the DAS[4] for rheumatoid arthritis (RA), the ASDAS is a composite index with continuous measurement properties. The development process resulted in four candidate ASDAS scores,[3] all of them fulfilling important aspects of truth, feasibility and discrimination.[3,5] The ASAS membership has selected the ASDAS with C-reactive protein (CRP) as the preferred version and with erythrocyte sedimentation rate (ESR) as the alternative version.[3]

In order to increase interpretability, a disease activity measure requires criteria for identifying 'disease activity states' (or 'status') and 'improvement' (or 'response criteria'). Improvement scores help to determine whether treatments really work, that is whether they actually produce clinically important improvement, allowing investigators, clinicians, regulators and patients to determine the efficacy (or lack thereof) of a given intervention and to communicate about response using the same metric.[6] Disease activity states measure clinical disease activity at specific timepoints. They are important for supporting decisions about entry into clinical trials, for supporting treatment changes and for defining therapeutic goals. Furthermore, in light of recent therapeutic advances and the increasing potential to improve the outcomes of patients with AS, the definition of criteria for disease states according to the ASDAS is highly relevant, as the prognosis may be different in patients depending on the disease activity states they attain, even if the same level of improvement is achieved. This observation highlights the importance of reporting disease activity states and not just absolute and categorical therapeutic responses, an important concept that has been clearly demonstrated in RA.[7]

Criteria for disease activity states and improvement scores are therefore important for use in clinical practice, observational studies and clinical trials and so far have not

been developed for the ASDAS. In the present study, we evaluated clinically relevant cut-off values for disease activity states and improvement scores using both forms of the ASDAS.

# PATIENTS AND METHODS

### ASDAS calculation

The ASDAS formulae[3] are as follows:

ASDAS-CRP (the preferred version):

0.12 × Back Pain + 0.06 × Duration of Morning Stiffness + 0.11 × Patient Global + 0.07 × Peripheral Pain/Swelling + 0.58×Ln(CRP+1)

ASDAS-ESR (the alternative version):

0.08 × Back Pain + 0.07 × Duration of Morning Stiffness + 0.11 × Patient Global + 0.09 × Peripheral Pain/Swelling + 0.29 × √(ESR)

CRP is in mg/litre, ESR is in mm/h; the range of other variables is from 0 to 10; Ln represents the natural logarithm; √ represents the square root.

### Nomenclature for ASDAS disease activity states and improvement scores

During the 2010 ASAS workshop in Berlin, Germany, upon presentation of results and discussion, four disease activity states and two improvement scores were chosen by consensus: (1) disease activity states: 'inactive disease', 'moderate disease activity', 'high disease activity' and 'very high disease activity'; and (2) improvement scores: 'minimal clinically important improvement' (MCII) and 'major improvement'.

### Study population used for the selection of cut-offs

For the selection of cut-offs we used data from the Norwegian disease modifying antirheumatic drug (NOR-DMARD) register[8,9] a Norwegian five-centre register that includes consecutive patients with AS (according to the treating doctor) starting a new conventional or biological DMARD regimen. Measures of disease activity and health status are assessed at baseline, 3, 6, 12 months and yearly thereafter. Patients from the NOR-DMARD register are an appropriate representation of patients with AS in general, as seen by rheumatologists in Norway. Of the patients from NOR-DMARD that we analysed, 69% were men, 90% were positive for human leucocyte antigen (HLA)-B27, the mean (SD) age was 43.3 (10.7) years and the mean disease duration since diagnosis

was 12.0 (10.6) years. Detailed characteristics of patients included in NOR-DMARD have been described previously.[8,9]



In order to have the best representation of the disease activity states being studied, 3-month data (n=331–336) were used to select the cut-off for 'inactive disease' and between 'moderate' and 'high disease activity', while baseline data (n=467–477) were only used to select the cut-off for 'very high disease activity'. The reason for this choice was because the large majority of patients from NOR-DMARD had (very) active disease at baseline (eg, none of the patients fulfilled ASAS partial remission criteria). Change scores between baseline and 3-month assessment (n=295) were used to select the cut-offs for improvement. The development of cut-offs was performed using ASDAS-CRP, the preferred ASDAS version.

### Study populations used for cross-validation of the cut-offs

Cross-validation was performed in NOR-DMARD (with an additional timepoint at 6 months) and in an 80% random sample of the Ankylosing Spondylitis Study for the Evaluation of Recombinant Infliximab Therapy (ASSERT) cohort (n=219–223).[10] In brief, ASSERT was a randomised 24-week placebo-controlled trial with infliximab that included patients with AS (according to the modified New York criteria[11]) with a Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)[12] and a spinal pain score ≥4 (range 0–10). The ASSERT population was typical of patients with moderate to severe AS. Of the patients from ASSERT that we analysed, 79% were men, 89% were positive for HLA-B27, the mean (SD) age was 39.3 (10.1) years and the mean disease duration was 10.6 (8.7) years. Detailed characteristics of patients in the ASSERT trial have been described previously.[10] For the validation we used baseline, 12-week and 24-week data.

The validation of the cut-offs was performed for ASDAS-CRP and ASDAS-ESR. Owing to the statistical approach used in the development of the ASDAS formulae,[3] it was expected that the cut-offs developed with ASDAS-CRP would also be applicable to ASDAS-ESR.

### Measurement instruments

Patient assessment of global disease activity and the six individual questions of the BASDAI were available in NOR-DMARD and ASSERT. The range of all scores is from 0 to 10. CRP (mg/litre) was also available in both databases, while ESR (mm/h) and physician's global assessment of disease activity were only available in NOR-DMARD. With these assessments, ASDAS-CRP could be calculated in both databases while ASDAS-ESR could only be calculated in NOR-DMARD.

In previous studies concerning the ASDAS,[3,5] no description has been given as to how values below the CRP threshold of detection should be handled. This has now been

studied and we recommend that in such cases half of the value of the threshold should be used (eg, if the limit of detection is 4 mg/litre, a value of 2 should be used). The use of the high sensitivity CRP assay is preferred.

The Bath Ankylosing Spondylitis Functional Index (BASFI)[13] was also available in both databases, allowing us to calculate ASAS partial remission and ASAS response criteria.[14 15] Moreover, having BASDAI total score available, we were also able to calculate response measures used for the evaluation of efficacy of anti-tumour necrosis factor (TNF) treatment in clinical practice, based on the BASDAI, that is the proportion of patients who had at least 2 units improvement (ΔBASDAI≥2) or at least 50% improvement (BASDAI50).

**Use of the receiver operating characteristic analysis for the selection of cut-offs in NOR-DMARD**

As there is no universal gold standard to assess disease activity in AS, we performed receiver operating characteristic (ROC) analysis against predefined external criteria considered to be representative of the various diseases activity states. Because ASDAS cut-offs should be representative of the perspectives of patients and doctors, we used the patient and physician global assessments at predefined levels (<1, <3 and >6 cm) as external constructs for 'inactive disease', to separate 'moderate' from 'high disease activity' and for 'very high disease activity', respectively. Additionally, for determining the cut-off for 'inactive disease' we also used ASAS partial remission as an external criterion (table 1).

One of the questions from ASAS members was about estimating the relationship between BASDAI and ASDAS as the BASDAI cut-off of 4 has been extensively used in trials with TNF blockers to determine 'high disease activity'. Therefore, we compared BASDAI (<3, <3.5 and <4 cm) with the cut-off between 'moderate' and 'high disease activity' (table 1).

Regarding improvement, the most frequently recommended external criterion for ROC analysis (an anchor-based approach) is the 'global rating of change' (GRC), a Likert-type scale scored for change by the patient.[16–18] In NOR-DMARD such a scale is available in the form of a unique question where patients score the change in their health status according to five categories: 'much better', 'better', 'unchanged', 'worse' and 'much worse'. For the ROC analysis, external anchors were constructed by dichotomising the rating scale for change in two different ways: a cut-off between 'much better/better' and 'unchanged/worse/much worse' in order to determine 'MCII', and a cut-off between 'much better' and 'better/unchanged/worse/much worse' to determine 'major improvement'. Moreover, we used the entire cohort in the ROC analysis, rather than just the two groups adjacent to the dichotomisation point because it has been shown that

this procedure maximises precision and yields a more logical estimate of the cut-offs.[19] The same principle was used in the ROC analysis for disease activity states.

We applied three methods of 'optimal' cut-off determination: (1) fixed 90% specificity, (2) the Youden index and (3) the closest point to (0,1), that is the point where the shoulder of the ROC curve is closest to the left upper corner of the graphic. The first method is particularly important in the clinical context (you try to avoid that patients in low/moderate disease activity are misclassified as inactive), while the last two methods provide the best balance between sensitivity and specificity.[20–22]



**Comparison of the cut-off for 'MCII' obtained by the ROC method with 'minimal detectable improvement' obtained by other methods**

The ROC method assesses which change on the measurement instrument corresponds with an important/meaningful change defined by the anchor, in this case the patient.[23] This is higher in hierarchy than 'minimal detectable improvement' based on measurement precision.[18] However, it is important to assure that the 'MCII' lies within boundaries that can be assessed beyond measurement error.[23] Therefore, we compared 'MCII' obtained by the ROC method with various methods of determining 'minimal detectable improvement' and used this to benchmark the choice of the cut-off value for 'MCII'.

Comparison was made with the 'mean change' (a less reliable anchor-based approach)[24] and several distribution based approaches: the 'Wyrwich standard error of measurement',[25] the 'Jacobson's reliable change index',[26] the '0.5*SD approach',[27] and the 'smallest detectable change approach'[28] (supplementary table 1).

**Cross-validation study**

Cross-validation was performed in NOR-DMARD and ASSERT for ASDAS-CRP and in NOR-DMARD for ASDAS-ESR. In order to allow comparisons between ASDAS-CRP and ASDAS-ESR, only patients with both values available were used for cross-validation in NOR-DMARD. However, including all patients with obtainable data for each ASDAS version (approximately 10% more patients) the results were similar (data not shown). Several cross-validation approaches were used:

1. Calculation of sensitivity and specificity of ASDAS cutoff values in comparison with several other criteria at different timepoints.
2. Assessment of the longitudinal distribution of patients over ASDAS disease activity states before and after start of treatment.
3. Mean values of BASDAI and ASDAS across the four ASDAS disease activity states.
4. Percentage of patients achieving ASDAS improvement criteria ('MCII' and 'major improvement') in comparison to other widely used improvement

criteria (ΔBASDAI≥2, BASDAI50, ASAS20 and ASAS40), 3 and 6 months after start of treatment.

5. In order to assess discriminative power, $\chi^2$ and p values were calculated for the differences between placebo and infliximab in ASSERT. SPSS V.17.0 (SPSS, Chicago, Illinois, USA) was used in all statistical analysis.

# RESULTS

**Selection of the optimal cut-offs for disease activity states and improvement scores**

The cut-offs for the various external criteria, according to fixed 90% specificity, Youden index and closest point to (0,1) are presented in table 1. The 90% specificity criterion was considered to be the most clinically relevant cut-off for 'inactive disease', to separate 'moderate' from 'high disease activity' and for improvement scores. In these cases, specificity is clinically more important in order to reduce the risk of misclassifying patients whose disease remains active (or who have not really improved) according to the external construct. Regarding the cut-off for 'very high disease activity', we considered that it would be better to have the best balance between sensitivity and specificity.

The definite choice for appropriate cut-offs was facilitated by consistent results across all external criteria (table 1). Such concordance between patient and physician global scores (and ASAS partial remission criteria, in the case of 'inactive disease') adds to the robustness of our results.

The three cut-offs for disease activity states selected after debate and voting by ASAS members were as follows: <1.3 between 'inactive disease' and 'moderate disease activity', <2.1 between 'moderate' and 'high disease activity' and >3.5 between 'high' and 'very high disease activity' (figure 1A). The cut-off between 'moderate' and 'high disease activity' (<2.1 units) corresponded to a BASDAI cut-off of <3.5 cm (table 1).

The cut-offs selected for improvements were: change of ≥1.1 units for 'MCII' and change of ≥2.0 units for 'major improvement' (figure 1B). Importantly, the cut-off for 'MCII' exceeded the 'minimal detectable improvement' based on measurement error, which ranged from 0.4 to 1.1 (supplementary table 1).

**Cross-validation results**

Regarding ASDAS-CRP, the cut-offs developed in NOR-DMARD at 3 months showed similar results in terms of sensitivity and specificity against the same (and other)

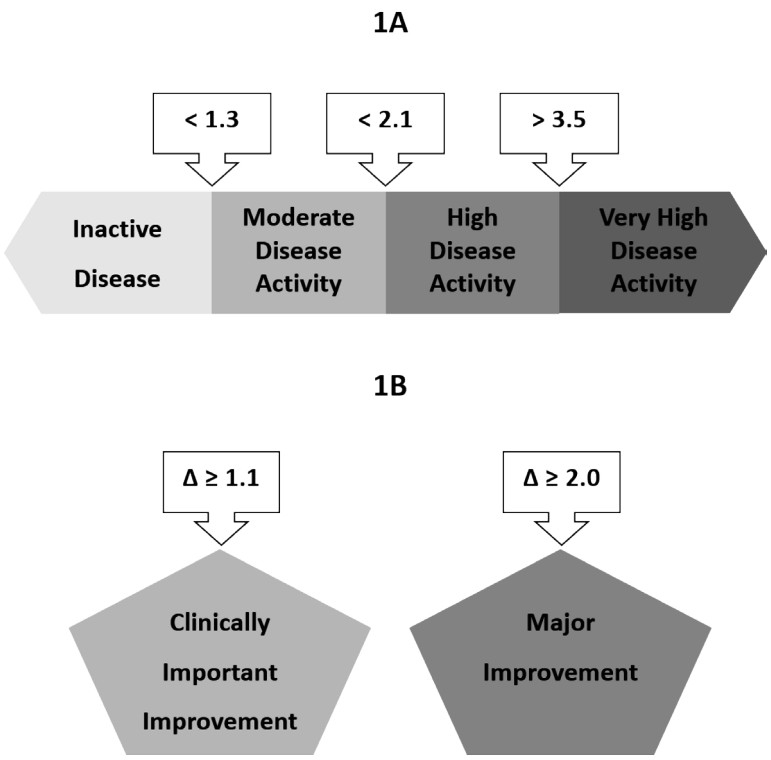


**Figure 1.** Selected cut-offs for (A) disease activity states and (B) improvement scores according to the Ankylosing Spondylitis Disease Activity Score (ASDAS). Every improvement beyond the 'minimal clinically important improvement' is a 'clinically important improvement'.

external constructs in NOR-DMARD at 6 months and in ASSERT at 3 and 6 months (table 2). Noticeably, results in ASSERT often surpassed the results in NOR-DMARD, yielding higher sensitivities (above 80%) while retaining the same level of specificity (approximately 90%). For the cut-off between 'high' and 'very high disease activity' (analysis only preformed at baseline) the slightly lower concordance probably reflects the higher subjectivity of the cut-off and a different selection criterion for the 'optimal' cut-off.

The longitudinal distribution of ASDAS-CRP disease activity states in both databases (table 3) showed a clinically and statistically significant shift of treated patients from higher disease activity states towards lower disease activity states. Interestingly, in the longitudinal analysis of ASSERT, the differences between the infliximab and placebo groups clearly discriminate between the two treatment arms: at 6-month follow-up 31.9% (infliximab) versus 0% (placebo) of the patients had 'inactive disease' ($p<0.001$), while 12.3% (infliximab) versus 53.6% (placebo) had 'very high disease activity' ($p<0.001$).

**Table 1.** ROC analysis: ASDAS cut-offs for disease activity states and improvement scores according to several external criteria and according to three different methods of 'optimal' cut-off determination

| ASDAS cut-offs and external criteria | n (P+N) | 90% SP (SE/SP) | Youden (SE/SP) | (0,1) (SE/SP) | AUC (95% CI) |
|---|---|---|---|---|---|
| Cut-off between 'inactive disease' and 'moderate disease activity': | | | | | |
| ASAS partial remission | 336 (74+262) | **<1.29** (64.9/90.1) | <1.60 (90.5/81.3) | <1.54 (89.2/82.4) | 0.91 (0.88 to 0.94) |
| Patient global <1 | 336 (77+259) | **<1.35** (75.3/90.0) | <1.52 (87.0/84.1) | <1.52 (87.0/84.2) | 0.91 (0.88 to 0.94) |
| Physician global <1 | 331 (113+218) | **<1.29** (44.2/89.9) | <1.95 (77.0/71.1) | <1.95 (77.0/71.1) | 0.79 (0.74 to 0.84) |
| Cut-off between 'moderate' and 'high disease activity': | | | | | |
| Patient global <3 | 336 (179+157) | **<2.08** (79.3/89.8) | <2.46 (93.3/77.7) | <2.36 (89.9/80.9) | 0.94 (0.92 to 0.96) |
| Physician global <3 | 331 (258+73) | **<2.14** (59.7/90.4) | <2.58 (75.2/80.8) | <2.58 (75.2/80.8) | 0.84 (0.79 to 0.89) |
| BASDAI <3 | 337 (154+183) | <1.94 (86.4/90.1) | <1.83 (85.1/92.3) | <2.05 (88.3/88.5) | 0.94 (0.92 to 0.97) |
| BASDAI <3.5 | 336 (181+155) | **<2.17** (85.1/89.7) | <2.14 (83.4/91.6) | <2.17 (85.1/89.7) | 0.95 (0.92 to 0.97) |
| BASDAI <4 | 336 (202+134) | **<2.24** (82.7/90.3) | <2.15 (79.7/94.8) | <2.34 (86.1/88.1) | 0.94 (0.92 to 0.97) |
| Cut-off between 'high' and 'very high disease activity': | | | | | |
| Patient global >6 | 477 (220+257) | >3.75 (61.4/89.9) | **>3.58** (71.8/84.4) | **>3.53** (73.6/82.5) | 0.86 (0.83 to 0.89) |
| Physician global >6 | 467 (55+412) | >4.62 (30.9/90.1) | **>3.33** (87.3/52.4) | **>3.50** (80.0/60.0) | 0.74 (0.67 to 0.81) |
| Cut-off for 'clinically important improvement': | | | | | |
| Better or much better* | 295 (209+86) | **≥1.12** (63.6/89.5) | ≥1.12 (63.6/89.5) | ≥0.68 (76.1/76.7) | 0.84 (0.79 to 0.88) |
| Cut-off for 'major improvement': | | | | | |
| Much better* | 295 (91+204) | **≥2.01** (56.0/90.2) | ≥1.32 (83.5/75.0) | ≥1.32 (83.5/75.0) | 0.85 (0.80 to 0.90) |

ASAS partial remission criteria are fulfilled if the value of the following four domains is below 2 (range 0–10): spinal pain, physical function measured by the BASFI, patient global assessment and inflammation measured as the mean of the last two BASDAI questions (severity and duration of morning stiffness).
*External criterion based on a unique NOR-DMARD question, in which patients rated the perceived change in their health status since the start of treatment on a five-point Likert-type scale (much better, better, unchanged, worse, much worse). The chosen cut-offs are highlighted in bold. Range of BASDAI and patient and physician global assessment is 0–10.
(0,1), cut-off according to the closest point to (0,1) criterion; 90% SP, cut-off according to the 90% specificity criterion; ASAS, Assessment of SpondyloArthritis international Society; ASDAS, Ankylosing Spondylitis Disease Activity Score; AUC, area under the curve; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; P+N, number of positive+negative results according to the external criterion; ROC, receiver operating characteristic; SE, sensitivity; SP, specificity; Youden, cut-off according to the Youden index criterion.

**Table 2.** Sensitivity and specificity of ASDAS cut-off values for disease activity states and levels of improvement against several external criteria in NOR-DMARD and ASSERT

| | | NOR-DMARD (ASDAS-CRP) | | | NOR-DMARD (ASDAS-ESR) | | | ASSERT (ASDAS-CRP) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **ASDAS cut-offs and external criterion** | **Timepoint** | **SE** | **SP** | **n** | **SE** | **SP** | **n** | **SE** | **SP** | **n** |
| ASDAS <1.3: | | | | | | | | | | |
| ASAS partial remission | 3 Months | 66.7 | 88.8 | 310 | 78.3 | 86.3 | 310 | 80.6 | 92.3 | 219 |
| | 6 Months | 60.0 | 91.2 | 192 | 82.2 | 91.2 | 192 | 87.2 | 90.0 | 219 |
| Patient global <1 | 3 Months | 69.9 | 90.7 | 310 | 71.2 | 85.2 | 310 | 82.1 | 89.6 | 220 |
| | 6 Months | 67.5 | 85.6 | 207 | 80.0 | 82.6 | 207 | 81.3 | 86.1 | 219 |
| Physician global <1 | 3 Months | 44.4 | 88.3 | 305 | 56.5 | 87.8 | 305 | – | – | – |
| | 6 Months | 71.9 | 85.6 | 199 | 84.4 | 82.6 | 199 | – | – | – |
| ASDAS <2.1: | | | | | | | | | | |
| Patient global <3 | 3 Months | 80.4 | 87.3 | 310 | 91.7 | 80.3 | 310 | 81.7 | 90.6 | 220 |
| | 6 Months | 83.6 | 89.7 | 207 | 93.6 | 85.6 | 207 | 87.2 | 83.5 | 220 |
| Physician global <3 | 3 Months | 59.3 | 88.7 | 305 | 70.0 | 83.9 | 305 | – | – | – |
| | 6 Months | 84.3 | 89.7 | 199 | 94.1 | 85.6 | 199 | – | – | – |
| ASDAS >3.5: | | | | | | | | | | |
| Patient global >6 | Baseline | 74.8 | 79.2 | 442 | 59.4 | 88.3 | 442 | 85.2 | 58.0 | 223 |
| Physician global >6 | Baseline | 79.2 | 59.4 | 432 | 67.9 | 71.5 | 432 | – | – | – |
| ΔASDAS ≥1.1: | | | | | | | | | | |
| Better or much better | 3 Months | 62.7 | 86.7 | 252 | 63.8 | 80.0 | 252 | – | – | – |
| | 6 Months | 61.7 | 96.8 | 164 | 60.2 | 90.3 | 164 | – | – | – |
| ASAS20 | 3 Months | 84.5 | 83.8 | 258 | 86.2 | 80.3 | 258 | 83.6 | 87.0 | 220 |
| | 6 Months | 83.3 | 79.3 | 165 | 85.9 | 81.6 | 165 | 84.1 | 90.3 | 219 |
| ΔASDAS ≥2.0: | | | | | | | | | | |
| Much better | 3 Months | 56.4 | 90.2 | 252 | 59.0 | 93.7 | 252 | – | – | – |
| | 6 Months | 53.8 | 90.2 | 164 | 50.0 | 91.1 | 164 | – | – | – |
| ASAS40 | 3 Months | 64.9 | 93.9 | 258 | 61.0 | 94.5 | 258 | 86.5 | 80.8 | 220 |
| | 6 Months | 60.4 | 93.8 | 165 | 54.7 | 93.8 | 165 | 83.7 | 90.2 | 219 |

ASAS20 and ASAS40 response criteria are based on four independent domains: spinal pain, physical function measured by the BASFI, patient global assessment and inflammation measured as the mean of the last two BASDAI questions (severity and duration of morning stiffness); ASAS20 treatment response is defined as improvement of ≥20% and ≥1 unit (range 0–10) in at least three of the four above domains, and no worsening of ≥20% and ≥1 unit in the remaining fourth domain; ASAS40 treatment response is defined as improvement of ≥40% and ≥2 units in at least three of the four above domains, and no worsening in the remaining fourth domain; ASAS partial remission criteria are fulfilled if the value for all four domains is below 2. Range of patient and physician global assessment is 0–10. ASAS, Assessment of SpondyloArthritis international Society; ASDAS, Ankylosing Spondylitis Disease Activity Score; ASSERT, Ankylosing Spondylitis Study for the Evaluation of Recombinant Infl iximab Therapy; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; NOR-DMARD, Norwegian register of disease modifying antirheumatic drugs.

**Table 3.** Longitudinal distribution of ASDAS disease activity states (%) in NOR-DMARD and ASSERT

| Timepoint | N | ASAS partial remission | ASDAS<1.3 | 1.3≤ASDAS<2.1 | 2.1≤ASDAS≤3.5 | ASDAS>3.5 |
|---|---|---|---|---|---|---|
| NOR-DMARD (ASDAS-CRP): | | | | | | |
| Baseline | 442 | 1.6 | 1.6 | 7.2 | 45.7 | 45.5 |
| 3 Months | 310 | 22.3 | 23.5 | 25.8 | 32.6 | 18.1 |
| 6 Months | 192 | 23.4 | 20.8 | 25.0 | 35.9 | 18.2 |
| NOR-DMARD (ASDAS-ESR): | | | | | | |
| Baseline | 442 | 1.6 | 2.0 | 11.3 | 53.2 | 33.5 |
| 3 Months | 310 | 22.3 | 28.1 | 30.6 | 31.0 | 10.3 |
| 6 Months | 192 | 23.4 | 26.0 | 27.1 | 33.9 | 13.0 |
| ASSERT (ASDAS-CRP): | | | | | | |
| Baseline | 223 | 0 | 0 | 1.3 | 29.1 | 69.5 |
| 3 Months | 219 | 16.4 | 19.6 | 20.5 | 38.8 | 21.0 |
| 6 Months | 219 | 17.8 | 23.7 | 20.5 | 32.9 | 22.8 |
| ASSERT (ASDAS-CRP) infliximab vs placebo ($\chi^2$, p value): | | | | | | |
| Baseline | 166 vs 57 | 0 vs 0 (NA) | 0 vs 0 (NA) | 1.2 vs 1.8 (0.1, 0.756) | 30.1 vs 26.3 (0.3, 0.586) | 68.7 vs 71.9 (0.2, 0.645) |
| 3 Months | 163 vs 56 | 21.5 vs 1.8 (11.8, 0.001) | 25.8 vs 1.8 (15.2, <0.001) | 26.4 vs 3.6 (13.3, <0.001) | 38.7 vs 39.3 (0.01, 0.933) | 9.2 vs 55.4 (53.5, <0.001) |
| 6 Months | 163 vs 56 | 23.3 vs 1.8 (13.2, <0.001) | 31.9 vs 0 (23.4, <0.001) | 23.3 vs 12.5 (3.0, 0.084) | 32.5 vs 33.9 (0.04, 0.846) | 12.3 vs 53.6 (40.6, <0.001) |

ASAS partial remission criteria are fulfilled if the value of the following four domains is below 2 (range 0–10): spinal pain, physical function measured by the BASFI, patient global assessment and inflammation measured as the mean of the last two BASDAI questions (severity and duration of morning stiffness).
ASAS, Assessment of SpondyloArthritis international Society; ASDAS, Ankylosing Spondylitis Disease Activity Score; ASSERT, Ankylosing Spondylitis Study for the Evaluation of Recombinant Infliximab Therapy; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; NOR-DMARD, Norwegian register of disease modifying antirheumatic drugs.

Moreover, 'inactive disease' according to the ASDAS had higher discriminatory capacity ($\chi^2$=23.4, p<0.001) than ASAS partial remission criteria ($\chi^2$=13.2, p<0.001).



Comparison of BASDAI and ASDAS mean values across the four ASDAS activity states during follow-up (table 4) showed that ASDAS disease activity states were in agreement with clinically relevant numerical differences in BASDAI mean values: BASDAI mean value for ASDAS 'inactive disease' ranged from 0.78 to 1.12, while for ASDAS 'very high disease activity' it ranged from 6.93 to 7.29 (scale 0–10).

Finally, in both databases, ASDAS 'MCII' (ΔASDAS≥1.1) was able to identify more patients with clinically meaningful improvement than the classical criteria: for example in ASSERT at 6-month follow-up, 57.5% of patients achieved ASDAS 'MCII', while 51.6%, 41.6% and 52.5% achieved ΔBASDAI≥2, BASDAI50 and ASAS20, respectively (table 5). ASDAS 'MCII' was also able to discriminate better between infliximab and placebo groups when compared to classical response criteria (higher $\chi^2$ values). Regarding ASDAS 'major improvement' (ΔASDAS≥2.0) it was often a more stringent criterion than ASAS40, supporting its validity as a measure of large improvement. Moreover, similarly to the 'MCII' cut-off, it showed a higher capacity to discriminate between active and placebo groups compared to usual response criteria (higher $\chi^2$ values).

Regarding ASDAS-ESR, overall the results of the cross-validation in NOR-DMARD were very similar to ASDAS-CRP (tables 2–5). No relevant differences were observed for 'improvement cut-offs', while regarding the cut-off values for disease activity states, ASDAS-ESR showed a trend to categorise slightly more patients in lower disease activity states compared to ASDAS-CRP (eg, in NOR-DMARD at 6 months 26.0% had 'inactive disease' according to ASDAS-ESR and 20.8% according to ASDAS-CRP) and slightly less patients in higher disease activity states (13.0% had 'very high disease activity' according to ASDAS-ESR and 18.2% according to ASDAS-CRP).

**Table 4.** BASDAI and ASDAS mean values and SD across the four ASDAS disease activity states in NOR-DMARD (n=310 at 3 months, n=192 at 6 months) and ASSERT (n=219 at 3 and 6 months)

| | NOR-DMARD (ASDAS-CRP) | | NOR-DMARD (ASDAS-ESR) | | ASSERT (ASDAS-CRP) | |
|---|---|---|---|---|---|---|
| **Disease activity states** | **BASDAI (mean±SD)** | **ASDAS (mean±SD)** | **BASDAI (mean±SD)** | **ASDAS (mean±SD)** | **BASDAI (mean±SD)** | **ASDAS (mean±SD)** |
| ASDAS <1.3 (3 months) | 1.09±0.87 | 0.94±0.26 | 1.12±0.79 | 0.92±0.21 | 0.97±0.65 | 0.94±0.22 |
| ASDAS <1.3 (6 months) | 1.01±0.67 | 0.90±0.29 | 1.04±0.73 | 0.91±0.22 | 0.78±0.60 | 0.95±0.20 |
| 1.3≤ASDAS<2.1 (3 months) | 2.17±1.26 | 1.62±0.22 | 2.60±1.30 | 1.66±0.24 | 2.38±0.99 | 1.65±0.23 |
| 1.3≤ASDAS<2.1 (6 months) | 2.37±1.23 | 1.64±0.22 | 2.78±1.15 | 1.67±0.20 | 2.53±1.06 | 1.70±0.24 |
| 2.1≤ASDAS≤3.5 (3 months) | 4.40±1.55 | 2.67±0.38 | 5.29±1.51 | 2.75±0.41 | 4.87±1.55 | 2.78±0.40 |
| 2.1≤ASDAS≤3.5 (6 months) | 4.59±1.73 | 2.75±0.41 | 5.23±1.62 | 2.73±0.42 | 4.92±1.40 | 2.80±0.39 |
| ASDAS>3.5 (3 months) | 6.93±1.33 | 4.12±0.63 | 7.29±1.34 | 4.31±0.62 | 7.07±1.57 | 4.31±0.57 |
| ASDAS>3.5 (6 months) | 7.04±1.42 | 4.23±0.58 | 7.24±1.53 | 4.20±0.52 | 7.19±1.31 | 4.33±0.63 |

ASAS, Assessment of SpondyloArthritis international Society; ASDAS, Ankylosing Spondylitis Disease Activity Score; ASSERT, Ankylosing Spondylitis Study for the Evaluation of Recombinant Infliximab Therapy; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; NOR-DMARD, Norwegian register of disease modifying antirheumatic drugs.

**Table 5.** Percentage of patients achieving ASDAS improvement criteria and classical improvement criteria in NOR-DMARD and ASSERT

| | NOR-DMARD (ASDAS-CRP) | | NOR-DMARD (ASDAS-ESR) | | ASSERT (ASDAS-CRP) | | ASSERT (ASDAS-CRP) infliximab vs placebo | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Improvement criterion** | **3 Months (n=258)** | **6 Months (n=165)** | **3 Months (n=258)** | **6 Months (n=165)** | **3 Months (n=220)** | **6 Months (n=219)** | **3 Months (n=164 vs 56)** | **$\chi^2$ (p value)** | **6 Months (n=163 vs 56)** | **$\chi^2$ (p value)** |
| ΔASDAS≥1.1 | 46.9 | 50.3 | 49.6 | 50.3 | 58.2 | 57.5 | 71.3 vs 19.6 | 45.9 (<0.001) | 69.3 vs 23.2 | 36.3 (<0.001) |
| ΔASDAS≥2.0 | 23.6 | 23.6 | 22.1 | 21.8 | 33.6 | 39.3 | 43.9 vs 3.6 | 30.4 (<0.001) | 50.9 vs 5.4 | 36.3 (<0.001) |
| ΔBASDAI≥2 | 43.0 | 43.6 | 43.0 | 43.6 | 50.9 | 51.6 | 60.4 vs 23.2 | 23.1 (<0.001) | 62.6 vs 19.6 | 30.8 (<0.001) |
| BASDAI50 | 36.8 | 39.4 | 36.8 | 39.4 | 40.5 | 41.6 | 50.6 vs 10.7 | 27.6 (<0.001) | 51.5 vs 12.5 | 26.1 (<0.001) |
| ASAS20 | 45.0 | 47.3 | 45.0 | 47.3 | 54.1 | 52.5 | 64.0 vs 25.0 | 25.6 (<0.001) | 63.2 vs 21.4 | 29.2 (<0.001) |
| ASAS40 | 29.8 | 32.1 | 29.8 | 32.1 | 41.8 | 38.8 | 50.6 vs 16.1 | 20.5 (<0.001) | 47.2 vs 14.3 | 19,1 (<0.001) |

ASAS20 and ASAS40 response criteria are based on four independent domains: spinal pain, physical function measured by the BASFI, patient global assessment and inflammation measured as the mean of the last two BASDAI questions (severity and duration of morning stiffness); ASAS20 treatment response is defined as improvement of ≥20% and ≥1 unit (range 0–10) in at least three of the four above domains, and no worsening of ≥20% and ≥1 unit in the remaining fourth domain; ASAS40 treatment response is defined as improvement of ≥40% and ≥2 units in at least three of the four above domains, and no worsening in the remaining fourth domain; ASAS partial remission criteria are fulfilled if the value for all four domains is below 2.

ASAS, Assessment of SpondyloArthritis international Society; ASDAS, Ankylosing Spondylitis Disease Activity Score; ASSERT, Ankylosing Spondylitis Study for the Evaluation of Recombinant Infliximab Therapy; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; NOR-DMARD, Norwegian register of disease modifying antirheumatic drugs.

# DISCUSSION

This study sought to determine cut-off values for disease activity states and improvement scores in AS based on the ASDAS. The definition of such criteria is of clinical and scientific importance.[6,7] We developed the cut-offs in a routine care population of patients with AS (NOR-DMARD) and validated them in the same population at a different timepoint and in a TNF blocker trial population (ASSERT). The fact that the cut-offs preformed at least as good in the trial population enhances their potential for application in both settings. Noticeably, the results of the cross-validation with ASDAS-CRP and ASDAS-ESR were very similar, supporting the use of the same cut-offs with both ASDAS versions.

The cut-offs were developed on clinical and statistical grounds and showed a remarkable consistence between the various external constructs that were used. Regarding improvement cut-offs, the availability of a GRC questionnaire in NOR-DMARD allowed us to use the most adequate gold standard for this purpose.[17,18,29] Importantly, the cut-off for 'MCII' was beyond borders of measurement error according to all tested methods.

ASDAS categories will facilitate studying the impact of disease activity states on prognosis. Furthermore, the cut-off for 'inactive disease' may be an important guideline for achieving a therapeutic aim. Compared to ASAS partial remission criteria, ASDAS 'inactive disease' has the advantage of being independent of BASFI: patients with a lot of structural damage that (as a consequence) have a high BASFI[30] may never achieve ASAS partial remission, while they may more easily achieve 'inactive disease'. In light of the results of the cross-validation, the new ASDAS-based improvement cut-offs may also facilitate the discrimination between treatment arms in clinical trials, and therefore result in smaller sample sizes.

The major limitation of our study is probably the lack of a universal and broadly accepted 'gold standard' for clinical disease activity in AS. However, we believe that the use of patient and physician global assessments as external constructs and their remarkable consistence for the selection of cut-offs overcomes this limitation. The use of arbitrary cut-offs for the external constructs may also be argued, but this was the only possible approach and the predefined cut-offs were discussed and accepted by ASAS members as representative of the disease activity states under study.

In summary, cut-off values for disease activity states and levels of improvement have been developed for the ASDAS. These cut-offs have proven to have external validity and a good performance in cross-validation. They have been endorsed by ASAS and are now ready to be used in clinical practice, observational studies and clinical trials.

# REFERENCES




1. Braun J, Sieper J. Ankylosing spondylitis. Lancet 2007;369:1379–90.
2. Bakker C, Boers M, van der Linden S. Measures to assess ankylosing spondylitis: taxonomy, review and recommendations. J Rheumatol 1993;20:1724–30.
3. 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.
4. van der Heijde DM, van 't Hof M, van Riel PL, et al. Validity of single variables and indices to measure disease activity in rheumatoid arthritis. J Rheumatol 1993;20:538–41.
5. van der Heijde D, Lie E, Kvien TK, et al. ASDAS, a highly discriminatory ASASendorsed disease activity score in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68:1811–18.
6. Singh JA, Solomon DH, Dougados M, et al. Development of classification and response criteria for rheumatic diseases. Arthritis Rheum 2006;55:348–52.
7. Aletaha D, Funovits J, Smolen JS. The importance of reporting disease activity states in rheumatoid arthritis clinical trials. Arthritis Rheum 2008;58:2622–31.
8. Kvien TK, Heiberg Lie E, et al. A Norwegian DMARD register: prescriptions of DMARDs and biological agents to patients with inflammatory rheumatic diseases. Clin Exp Rheumatol 2005;23(5 Suppl 39):S188–94.
9. Heiberg MS, Nordvåg BY, Mikkelsen K, et al. The comparative effectiveness of tumor necrosis factor-blocking agents in patients with rheumatoid arthritis and patients with ankylosing spondylitis: a six-month, longitudinal, observational, multicenter study. Arthritis Rheum 2005;52:2506–12.
10. van der Heijde D, Dijkmans B, Geusens P, et al. Efficacy and safety of infliximab in patients with ankylosing spondylitis: results of a randomized, placebo-controlled trial (ASSERT). Arthritis Rheum 2005;52:582–91.
11. 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.
12. 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.
13. Calin A, Garrett S, Whitelock H, 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.
14. Anderson JJ, Baron G, van der Heijde D, et al. Ankylosing spondylitis assessment group preliminary definition of short-term improvement in ankylosing spondylitis. Arthritis Rheum 2001;44:1876–86 .
15. Brandt J, Listing J, Sieper J, et al. Development and preselection of criteria for short term improvement after anti-TNF alpha treatment in ankylosing spondylitis. Ann Rheum Dis 2004;63:1438–44.
16. Wyrwich KW, Wolinsky FD. Identifying meaningful intra-individual change standards for health-related quality of life measures. J Eval Clin Pract 2000;6:39–49.
17. Turner D, Schünemann HJ, Griffi th LE, et al. The minimal detectable change cannot reliably replace the minimal important difference. J Clin Epidemiol 2010;63:28–36.
18. Tubach F, Ravaud P, Beaton D, et al. Minimal clinically important improvement and patient acceptable symptom state for subjective outcome measures in rheumatic disorders. J Rheumatol 2007;34:1188–93.
19. Turner D, Schünemann HJ, Griffith LE, et al. Using the entire cohort in the receiver operating characteristic analysis maximizes precision of the minimal important difference. J Clin Epidemiol 2009;62:374–9.
20. Coffin M, Sukhatme S. Receiver operating characteristic studies and measurement errors. Biometrics 1997;53:823–37.
21. Youden WJ. Index for rating diagnostic tests. Cancer 1950;3:32–5.
22. Perkins NJ, Schisterman EF. The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. Am J Epidemiol 2006;163:670–5.
23. de Vet HC, Terwee CB, Ostelo RW, et al. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. Health Qual Life Outcomes 2006;4:54.
24. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials 1989;10:407–15.
25. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEMbased criterion for identifying meaningful intra-individual changes in health-related quality of life. J Clin Epidemiol 1999;52:861–73.
26. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult

Clin Psychol 1991;59:12–19.

27. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care 2003;41:582–92.
28. Bruynesteyn K, Boers M, Kostense P, et al. Deciding on progression of joint damage in paired films of individual patients: smallest detectable difference or change. Ann Rheum Dis 2005;64:179–82.
29. de Vet HC, Terluin B, Knol DL, et al. Three ways to quantify uncertainty in individually applied "minimally important change" values. J Clin Epidemiol 2010;63:37–45.
30. Landewé R, Dougados M, Mielants H, et al. Physical function in ankylosing spondylitis is independently determined by both disease activity and radiographic damage of the spine. Ann Rheum Dis 2009;68:863–7.

# SUPPLEMENTARY MATERIAL



**Supplementary table 1.** ASDAS minimal detectable improvement

| Method for calculating MDI | Measurement error |
|---|---|
| Mean change of stable patients between 0-3 months | 1.05 |
| Wyrwich SEM | 0.41 |
| Jacobson's RCI | 1.13 |
| 0.5*SD of change between 0-3 months | 0.62 |
| SDC of stable patients between 0-3 months | 1.06 |

MDI, minimal detectable improvement; ASDAS, Ankylosing Spondylitis Disease Activity Score; SEM, standard error of measurement; RCI, reliable change index; SD, standard deviation; SDC, smallest detectable change. Mean change: the minimal detectable improvement (MDI) is the mean Δscore of patients who had small improvement ('better' on the global rating of change). Wyrwich SEM: MDI= $SD_{BL}$ x (√[1-r]). Jacobson's RCI: MDI= 1.96 x $SD_{BL}$ x (√(2 x [1-r])). 0.5 SD approach: the MDC is 0.5 SD of the Δscore of the instrument between 2 time-points. SDC approach: MDI= 1.96 x (SD of Δscore in 'unchanged' patients between 2 time-points)/√2. For the Wyrwich SEM, the test-retest intraclass correlation coefficient of stable patients was used for 'r'; BL, baseline.