

Just released from the ASAS factory! First steps towards a disease activity score for ankylosing spondylitis

Maarten Boers

The burden of disease in ankylosing spondylitis (AS) has been hard to grasp for doctors not directly involved in the care of these patients. With inflammation primarily affecting the axial skeleton, the outward manifestations are limited until late in the disease, when irreversible stiffening (ankylosis) and sometimes deformation progressively limit physical function and lead to disability. Society has little patience for patients with AS where so many people have complaints in the lower back area, and the insidious nature of AS allows patients to stay active participants in society for a very long time. However, the rapid uptake of biological agents, the first truly effective treatment beyond non-steroidal anti-inflammatory drugs revealed a large unmet need to alleviate the suffering caused by pain and stiffness.

The Outcome Measures in Rheumatology (OMERACT) initiative started in 1992 with a consensus conference on rheumatoid arthritis (RA) outcome measures, but very early on scientists involved in the care of patients with AS spun off as an independent group to develop a core set of domains to measure in AS. This spin off was called ASAS (Assessment of SpondyloArthritis international Society), and its results were first presented at OMERACT 4. Since then, the group has comprised a tightly integrated and highly motivated group of clinical researchers, independently producing first-class material in the area of outcome measurement, whilst remaining closely linked to OMERACT, both through adoption of OMERACT methodology and by bringing ASAS results to the OMERACT meetings for endorsement. It is no exaggeration to say that the development of AS as an indication for biological treatment was greatly

speeded up by the work of ASAS on measurement methodology.

Science is never done, so ASAS justifiably went about improving what is available: in this issue of the *Annals* ASAS presents the first step in the development of a disease activity score (*see page 18*).¹ Readers probably need no reminder that the senior author on this paper, Dr van der Heijde was the first author on a paper in this journal in 1990 describing the first step in what was to become the disease activity score (DAS) in RA.² The construct of disease activity remains a difficult concept to get a grip on in RA, and perhaps this is even more so in AS. Reasons for this include: (a) a lack of understanding of the pathophysiology of the disease, forcing us to treat the phenotype rather than the cause; (b) lack of measures that comfortably pass the OMERACT filter of truth, discrimination and feasibility³; (c) lack of good treatment options that allow validation of measures. In the case of AS, despite the availability of measures and one index (composite of measures) that had passed the filter (and allowed the trials of biological agents to proceed as they did), members of ASAS felt that there was room for improvement. This led to the initiative to derive a new index by statistical means, closely mimicking the development of the DAS.

The article contains a very clear description of the methodology, and is only briefly summarised here to highlight important issues. The process started with a Delphi exercise among experts to collect all potentially relevant measures, followed by a meeting to resolve outstanding problems. This step is highly important to ensure that no relevant measures are missed (content validity), but also to increase the chance of buy-in to the methodology and adoption of the final index even if it does not contain a measure favoured by some experts. It was decided not to include a measure of function in the index as a matter of principle: function can also irreversibly

be affected by damage. In RA, this is one of the main differences between the DAS and the WHO/ILAR core set^{4,5}: the exclusion or inclusion of a function measure. In RA, the Health Assessment Questionnaire (HAQ) is among the most discriminative measures for disease activity, despite some loss of sensitivity in very damaged patients,⁶ and I personally think it was a mistake not to include a measure of function, at least in the testing phase. Another problem with the approach taken, as acknowledged by the authors, is the lack of patient input in this stage. Finally, experts suggested also testing indices not containing acute phase reactants or patient global assessment to increase feasibility.

The second step was to collect a series of patients and then to distinguish between those with high and low disease activity. In this step lies the eternal problem of construct validation: it is inherently circular because one is building and validating a new (imperfect) measure by comparing it with another (imperfect) measure. The authors chose as comparison the decision taken by the treating doctor whether or not treatment with tumour necrosis factor (TNF) antagonists should be started. This approach has high face validity because presumably such decisions are primarily taken on the basis of an implicit assessment of disease activity. A problem recognised by the authors was that the decision was theoretical: the participating doctors had no experience with this treatment at the time the question was asked, and treatment was not really started. This in contrast with the RA-DAS studies where data on actual treatment decisions were used. Notably, ASAS was able to do a huge survey of 1200 patients through their network of participating rheumatologists from many different countries.

The third step was to decrease redundancy of the information in the different disease activity measures through factor analysis, followed by discriminant analysis to derive a formula that optimally distinguishes between the groups (with high versus low disease activity). This formula is then back translated to derive a linear formula with the most important single measures resulting in the candidate AS-DAS. A potential weakness of the approach is that only patients with complete data can be included, substantially reducing the dataset to about 700 patients. Also, the best formula correctly classified only 72% of patients as having either high or low disease activity, a

Correspondence to: Professor M Boers, Department of Epidemiology and Biostatistics, VU University Medical Centre, PK 6Z 185, PO Box 7057, 1007 MB Amsterdam, The Netherlands; keb.info@vumc.nl

modest improvement over the 50/50 a priori distribution (the equivalent of tossing a coin).

The final step was to validate the AS-DAS in a new set of patients, conveniently available in the form of the OASIS (Outcome in Ankylosing Spondylitis International Study) cohort. Here, the authors distinguished two groups of patients with high and low disease activity based on the physician global assessment. The result of this impressive exercise is a set of no less than four candidate indices ready for further validation. Topics to be examined include sensitivity to change, and especially discrimination between a group that changes little and one that changes a lot (as in the primary analysis of a placebo-controlled clinical trial). However, it is highly likely that any of the four indices will prove better than any one of the single measures currently used. This is inherent in the way the indices have been constructed, maximising the available information (the signal) and reducing the random error associated with measurement (the noise). In fact, an abstract presented at this year's EULAR meeting suggests such an outcome, with little difference between each of the four candidate indices applied to data of the NORDMARD database.⁷ The Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), the venerable index used to date was beaten, but not soundly, confirming its high value despite a much less formal development path. And all indices correlated well with

patient global assessment (less well with physician global assessment, surprisingly).

So are all problems of AS disease activity measurement now settled? Well, not entirely. Preferably we would like to end up with only one index. In my view, the concerns around the inclusion of a function measure have not been resolved. And none of the candidates are elegant: all include the same clunky mix of log or root transformed factors, and coefficients with no less than three numerals behind the decimal. Given that the best discrimination is only 72%, perhaps other candidates should be tried that allow a more natural derivation without calculators or computers. This has been the lingering problem with the DAS in its many forms in RA, and has led to the development of yet another set of indices by our Austrian colleagues, the Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI).⁸ These indices are simple sums of the component measures and yet retain much, if not all of the power of the DAS. Thus, before things are fixed in stone, the ASAS group should seriously consider a simpler form of the index. And ASAS should issue a strong caveat that much experience in practice is needed before any index can even be considered for use to drive decisions in individual patient care.

But I would not like to end on a grumpy note. As a member of the OMERACT Executive I am delighted that ASAS remains closely linked to OMERACT. The authors and the ASAS

group are to be congratulated as they go from strength to strength, building the methodology that enables the research for better treatment in AS.

Competing interests: None.

Ann Rheum Dis 2009;**68**:1–2.
doi:10.1136/ard.2008.097774

REFERENCES

1. **Lukas C**, Landewé R, Sieper J, Dougados M, Davis J, Braun J, *et al*. Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. *Ann Rheum Dis* 2009;**68**:18–24.
2. **van der Heijde DM**, van't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, *et al*. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Ann Rheum Dis* 1990;**49**:916–20.
3. **Boers M**, Brooks P, Strand V, Tugwell P. The OMERACT Filter for outcome measures in rheumatology. *J Rheumatol* 1998;**25**:198–9.
4. **Felson DT**, Anderson J, Boers M, Bombardier C, Chernoff M, Fried B, *et al*. American College of Rheumatology preliminary core set of disease activity measures for use in rheumatoid arthritis clinical trials. *Arthritis Rheum* 1993;**36**:729–40.
5. **Boers M**, Tugwell P, Felson DT, van Riel PL, Kirwan JR, Edmonds JP, *et al*. World Health Organization and International League of Associations for Rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. *J Rheumatol* 1994;**21**:86–9.
6. **Aletaha D**, Strand V, Smolen JS, Ward MM. Treatment-related improvement in physical function varies with duration of rheumatoid arthritis: a pooled analysis of clinical trial results. *Ann Rheum Dis* 2008;**67**:238–43.
7. **Van Der Heijde D**, Landewé R, Lie E, Kvien TK. Cross-validation of 4 disease activity scores for ankylosing spondylitis (ASDAS) in the NORDMARD database [abstract]. *Ann Rheum Dis* 2008;**67**(Suppl II):58.
8. **Aletaha D**, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol* 2005;**23**:S100–8.

Quality & Safety in Health Care

Quality & Safety in Health Care is a leading international peer-review journal in the growing area of quality and safety improvement. It provides essential information for those wanting to reduce harm and improve patient safety and the quality of care. The journal reports and reflects research, improvement initiatives and viewpoints and other discursive papers relevant to these crucial aims with contributions from researchers, clinical professionals and managers and experts in organisational development and behaviour.

qshc.bmj.com

Quality
& Safety
in Health Care