

Commentary

The validity of a rheumatoid arthritis medical records-based index of severity compared with the DAS28

Robert Landewé and Désirée van der Heijde

University Hospital Maastricht, Department of Internal Medicine/Rheumatology, Maastricht, The Netherlands

Corresponding author: Robert BM Landewé, rlan@sint.azm.nl

Published: 31 March 2006

This article is online at <http://arthritis-research.com/content/8/3/107>

© 2006 BioMed Central Ltd

Arthritis Research & Therapy 2006, **8**:107 (doi:10.1186/ar1937)

See related research by Sato *et al.* in this issue [<http://arthritis-research.com/content/8/3/R57>]

Abstract

Outcome measures play an extremely important role in clinical trials and observational research. Outcome measures for rheumatoid arthritis cover a whole array of domains, ranging from measures describing the inflammatory process to measures describing the ultimate consequences of long-term disease, such as joint damage, physical function and quality of life. There is a scientific need to be able to quantify what is called the 'severity of rheumatoid arthritis', so that patients with rheumatoid arthritis can be clustered according to their propensity to develop an unfavourable outcome. It is a challenge to find an appropriate measure for severity. One attempt has been the development of the Rheumatoid Arthritis Medical Record-Based Index of Severity. This commentary elaborates on how such a measure of severity should be validated to determine whether it is appropriate for practical use.

In the present issue of *Arthritis Research and Therapy*, Sato and colleagues report on their effort to determine the convergent validity of the Rheumatoid Arthritis Medical Record-Based Index of Severity (RARBIS), a newly developed measure of severity in rheumatoid arthritis (RA) [1]. The authors claim moderate convergent validity with the disease activity score based on 28 joints (DAS28), and they propose the RARBIS as a tool to adjust for confounding by indication, a treatment bias that is introduced in observational studies where the severity of the disease determines the intensity of the treatment [2].

In the present commentary we shall discuss two issues. The first addresses the question of when a particular measure can be called 'validated'. The second issue discussed argues whether a statistical correlation of the RARBIS with the DAS28 adds to the validity of the RARBIS as a measure of severity in RA.

Validation of outcome assessments is a vague playing field where clear guidance is lacking. People hardly, if ever, speak the same language if they claim that a measure is validated, and definitions of subcategories of validation show wide overlap. A good example is the confusion about content validity versus construct validity. The former refers to the user's perception of the content of an instrument ('does it make sense?'), the latter pertains to the underlying construct (DAS28 and its association with inflammation in the joint), and very often both are used to describe the same thing.

The Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) initiative has provided a useful alternative to avoid such confusion [3]. The three-step OMERACT filter, which is OMERACT's framework for validation of outcome measures, prescribes that a measure should be truthful, discriminatory and feasible before it should be used.

Truth here refers to scientific evidence that the RARBIS really reflects what it is intended to measure – the severity of RA. It requires some association with other pivotal outcomes in RA, and it should be recognisable to experts in the field.

Discrimination incorporates the important domain of reliability: Does the RARBIS arrive at the same score when used by different assessors, or used repetitively under unchanged conditions? Discrimination also implies that the RARBIS can distinguish RA patients with mild RA from those with severe RA. Discrimination also questions whether the RARBIS score, if applied as an outcome measure in clinical trials, indeed shows change when the circumstances are really changing; for example, by the impact of effective therapy (sensitivity to change)?

Feasibility, finally, pertains to the ease with which the RARBIS can be elicited and calculated in practice. Time and costs are issues of interest here.

It is important to state that the OMERACT filter is not an obligatory framework, and that it does not prioritise research efforts. Priority depends on the context in which the outcome measure will be used, and it is the context rather than the OMERACT filter that should prescribe the path of validation. If the RARBIS will not be used as an outcome measure in clinical trials, it does not make sense to test the sensitivity to change. But if the RARBIS is used in observational studies to adjust for potential confounding by indication, it is critical to investigate the reliability of a RARBIS score in individual patients. Validation of outcome measures is not a single project. It implies an array of different studies in different contexts, all aiming at different aspects of validation. Validation is a continuous scientific process, often taking years, and is never complete.

This brings us to the second issue: How important is convergent validity with the DAS28 for the RARBIS? It makes sense to distinguish process variables from outcome variables. Outcome variables measure the consequences of disease. In chronic diseases such as RA the outcome is assessable only after many years. Process variables measure the intensity of the disease process, which, if sustained for a sufficiently long time, ultimately leads to irreversible consequences (outcome). Some process variables have predictive validity, which means that they can predict a certain outcome. A good example is the DAS28, which can predict radiographic progression over time. The question may arise whether the RARBIS is an outcome or a process variable. Looking at the content of the index, it seems as if the RARBIS combines domains of outcome (surgery, radiographic damage), domains reflecting the disease process (clinical status, acute phase reactants), and variables with predictive value (rheumatoid factor). It is not our intention to criticise the content of the RARBIS here, but from a conceptual point of view one may seriously question the composition of domains in relation to its main intention – the quantification of severity. This depends strongly on the concept of severity that one adheres to: Does severity only refer to irreversible aspects of the disease, or do principally reversible aspects of the disease also belong to the concept of severity?

What about convergent validity? The DAS28 is a process measure of disease activity, incorporating clinical status (joint counts and global health) and acute phase reactants [4]. Disease activity is associated with radiographic damage [5] and with surgery [6], both components of the RARBIS. Acute phase reactants are also an independent component of the RARBIS. It is therefore not surprising at all that the RARBIS correlates to some extent with the DAS28, and this correlation *per se* does not add to the validity of the RARBIS as a measure of severity in RA in our opinion. Reweighting

the subscales of the RARBIS with the aim to improve the correlation with the DAS28 will only lead to an overvaluation of the domains related to disease activity in an index intended to assess severity. The RARBIS does not improve conceptually by such a statistical effort. Only substantial – and not statistical – arguments that favour an increased weight of disease activity (clinical status) in a severity index could do that.

Fundamental considerations rather than statistical inference should guide the validation of outcome measures. In our opinion, the value of the RARBIS in truly determining the severity of RA in individual patients depends mainly on issues other than convergent validity with the DAS28.

Competing interests

The authors declare that they have no competing interests.

References

1. Sato M, Schneeweiss S, Scranton R, Katz JN, Weinblatt ME, Avorn J, Ting G, Shadick NA, Solomon DH: **The validity of a rheumatoid arthritis medical records-based index of severity compared with the DAS28.** *Arthritis Res Ther* 2006, **8**:R57.
2. Landewe RB: **The benefits of early treatment in rheumatoid arthritis: confounding by indication, and the issue of timing.** *Arthritis Rheum* 2003, **48**:1-5.
3. Boers M, Brooks P, Strand CV, Tugwell P: **The OMERACT filter for outcome measures in rheumatology [editorial].** *J Rheumatol* 1998, **25**:198-199.
4. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL: **Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis.** *Arthritis Rheum* 1995, **38**:44-48.
5. Welsing PM, Landewe RB, van Riel PL, Boers M, van Gestel AM, van der Linden S, Swinkels HL, van der Heide DM: **The relationship between disease activity and radiologic progression in patients with rheumatoid arthritis: a longitudinal analysis.** *Arthritis Rheum* 2004, **50**:2082-2093.
6. Wolfe F, Zwiilich SH: **The long-term outcomes of rheumatoid arthritis: a 23-year prospective, longitudinal study of total joint replacement and its predictors in 1,600 patients with rheumatoid arthritis.** *Arthritis Rheum* 1998, **41**:1072-1082.