

Commentary

Pooled indices to measure rheumatoid arthritis activity: a good reflection of the physician's mind?

Daniel Aletaha

National Institutes of Arthritis, Musculoskeletal, and Skin Diseases, National Institutes of Health, Bethesda, Maryland, USA

Corresponding author: Daniel Aletaha, aletahad@mail.nih.gov

Published: 14 December 2005

This article is online at <http://arthritis-research.com/content/8/1/102>

© 2005 BioMed Central Ltd

Arthritis Research & Therapy 2006, **8**:102 (doi:10.1186/ar1870)

See related research by Vander Cruyssen *et al.* in issue 7.5 [<http://arthritis-research.com/content/7/5/R1063>]

Abstract

Several pooled indices for the assessment of rheumatoid arthritis disease activity are available to rheumatologists. Face and criterion validity of these instruments can be assessed by determining the association of their measurements with opinions of physicians. Several confounding aspects must be considered in such analyses, especially blinding of the person(s) making the decisions to the instruments being studied and to the objective of the study in general. From several studies in the literature, there is currently no evidence that any one of the available composite indices is better or worse than any other. The choice of index in clinical practice should ideally be based on practical considerations related to the needs of the rheumatologist in the respective health care setting.

Several pooled indices are available to measure rheumatoid arthritis (RA) activity on a continuous scale [1]. These include the Disease Activity Score (DAS), the DAS based on 28 joint counts (DAS28), the Simplified Disease Activity Index (SDAI), and the Clinical Disease Activity Index (CDAI). These indices are essentially based on the same attributes of RA: joint counts, the patient's evaluation of RA activity, and acute-phase reactants. The SDAI and CDAI also include the physician's assessment of RA activity, and do not require transformation or weighting of the individual components or the use of a calculator. In addition, the CDAI is the only index that does not include a measure of acute-phase response. All of these individual variables that are pooled in the various indices have face validity in the context of measuring RA activity. In other words, each variable, if available, is likely to be considered in the implicit clinical assessment of RA activity by the clinician. When measures are integrated to yield pooled indices, the question slightly shifts toward how much the composite number produced by the index relates to a clinician's intuitive integration of the available measures of

disease activity. In other words, is the obtained number higher in those patients who physicians consider to have more active disease, and is it lower in those considered to have less active disease?

In a recent issue of *Arthritis Research & Therapy*, Vander Cruyssen and colleagues [2] aimed to compare this aspect of validity for the various disease activity indices. The results from their study suggest that the DAS28 is the best determinant of physician opinion, based on each physician's decision to increase or not increase the dose of infliximab in patients with RA in a real life setting. This method is very typical and was used in the past, for example, to derive the original DAS and the DAS28 [3,4].

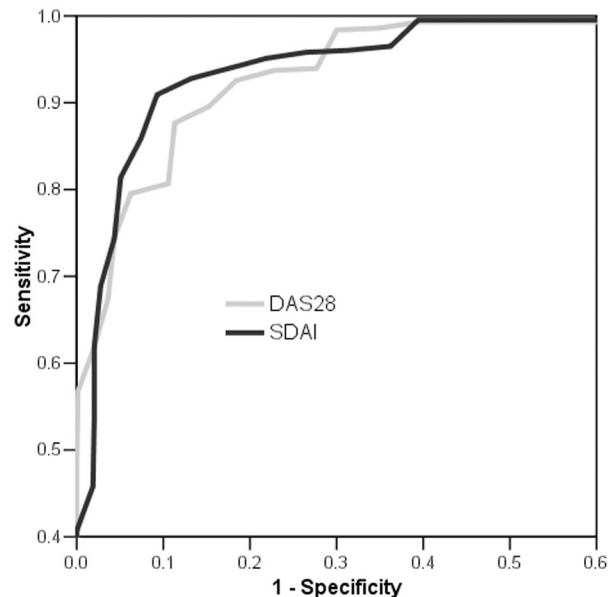
Several issues need consideration when analyses are based on physician opinion, and many of these issues are difficult to implement in a study setting. First, ideally physicians should not be aware that their clinical decision is part of the investigation. This issue can be regarded as an analogy to the classical epidemiologic problem of 'bias by observation': the fact that physicians are aware that their behavior is being observed is likely to make them more cautious and more considerate in their decisions as compared with their usual 'protected' clinical environment. Second, the physician's decision – as the 'gold standard' – should not be influenced by variables and measures that are planned to be used as independent predictors of this gold standard decision in subsequent statistical analyses. For example, one would expect that even a relatively unimportant measure would have a greater association with the clinician's decisions if it was the only measure available to that physician, and was thus the only measure objectively informing the decision.

In the study conducted by Vander Cruyssen and coworkers [2], this second issue is a potential concern. From the report it is unclear what was the basis for the treating rheumatologists' decision to increase the infliximab dose. The authors indicated that the composite scores 'were calculated after data collection so that the treating rheumatologist was unaware of the exact values of those composite scores'. However, in a previous report on this cohort [5], the authors stated that 'the ACR [American College of Rheumatology] response criteria and the DAS28 score were evaluated at the same time points before the infliximab infusion'. This statement indicates that calculation of the DAS28 was required per study protocol, and that this preceded the decision regarding whether or not to increase the dose of infliximab. Because other variants of the DAS28, SDAI, or CDAI were not calculated during the study, this can potentially magnify the ability of the DAS28 to predict a physician's decision in comparison with these other indices. On the other hand, if the physicians were blinded to the evaluation of ACR response and the DAS28, then it is unclear why the authors emphasized that this was done 'before the infliximab infusion'. Importantly, however, the authors revealed a high correlation coefficient among all investigated indices ($r = 0.9$ or higher) and mentioned that 'all those alternative scores perform similarly or slightly worse than the original DAS28', which is surprisingly good performance if only the DAS28 was calculated during the study.

Another issue was emphasized in the commentary by van Riel and Franssen [6]; more than 50% of patients in whom the infliximab dose was not increased had a DAS28 score above 3.2, which indicates moderate or high disease activity. This could mean that the treating rheumatologists either neglected to treat patients with significant disease activity more aggressively or that there is an inconsistency between the clinical characteristics of the patients and the DAS28 score. This issue of potentially poor sensitivity of the DAS28 criteria in identifying patients with moderate disease activity is not discussed in the report by Vander Cruyssen and coworkers. In fact, it supplements findings that the weighting of variables in the DAS28 may misrepresent the actual disease activity, especially in the lower disease activity ranges, because it weights erythrocyte sedimentation rate and tender joint counts quite strongly [7,8]. Taken together these issues would reduce the 'face validity' of the construction of the index.

We recently conducted a study to derive cutoff values for the DAS28 and the SDAI [7]. This study was based on the ratings of 35 expert rheumatologists. We re-analyzed these data to correlate SDAI and DAS28 scores of 32 paper patients (i.e. disease activity profiles of real RA patients) with the gold standard of physician's judgment, and performed a receiver operating characteristic (ROC) analysis (Fig. 1). The gold standard for this analysis was the physicians' judgment of moderate or high disease activity. We found that the (untransformed) SDAI exhibited an area under the ROC curve

Figure 1



Performance of the SDAI and the DAS28. Receiver operating characteristic curve analysis of the performance of the SDAI and the DAS28, using expert opinion on patient profiles as the gold standard. The experts rated whether moderate or high disease activity was present or not. DAS28, Disease Activity Score based on 28-joint evaluation; SDAI, Simplified Disease Activity Index.

of 0.96 (95% confidence interval 0.95–0.97); the DAS28 was similar, at 0.95 (95% confidence interval 0.94–0.96). The sensitivity at 95% specificity (as analyzed in the study by Vander Cruyssen and coworkers) was 80.5% for the SDAI and 76.2% for the DAS28. Likewise, a study conducted by Soubrier and coworkers [9] found an area under the ROC curve of 0.91 for the SDAI and of 0.86 for the DAS28, using a rheumatologist's decision to start a new disease-modifying antirheumatic drug as the gold standard.

Taking together the data from the study by Vander Cruyssen and coworkers [2] and others, we conclude that the DAS28 and SDAI exhibit similar face validity/criterion validity, with the potential exception of the very low disease activity ranges [6,7]. Arguments on validity that are based on minimal differences can be confusing for the rheumatologic community, especially when these differences show opposite directions in different studies. At the current time, there is no evidence that one index is better or worse than another.

The SDAI and the CDAI were not developed to oppose the DAS or DAS28 but to provide rheumatologists with a simple tool that can be calculated on the spot (CDAI) and without the need for a calculator (SDAI and CDAI). The fact that all of these scores have similar validity increases the choice of instruments available to physicians, allowing them to pick the index that best fits their practical needs and their environmental setting.

Competing interests

The author(s) declare that they have no competing interests.

References

1. Aletaha D, Smolen JS: **The definition and measurement of disease modification in inflammatory rheumatic diseases.** *Rheum Dis Clin N Am* 2005:in press.
2. Vander Cruyssen B, Van Looy S, Wyns B, Westhovens R, Durez P, Van den Bosch F, Veys EM, Mielants H, De Clerck L, Peretz A, et al.: **DAS28 best reflects the physician's clinical judgment of response to infliximab therapy in rheumatoid arthritis patients: validation of the DAS28 score in patients under infliximab treatment.** *Arthritis Res Ther* 2005, **7**:R1063-R1071.
3. van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, van Rijswijk MH, van de Putte LB: **Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score.** *Ann Rheum Dis* 1990, **49**:916-920.
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. Durez P, Van den Bosch F, Corluy L, Veys EM, De Clerck L, Peretz A, Malaise M, Devogelaer JP, Vastesaeger N, Geldhof A, et al.: **A dose adjustment in patients with rheumatoid arthritis not optimally responding to a standard dose of infliximab of 3 mg/kg every 8 weeks can be effective: a Belgian prospective study.** *Rheumatology (Oxford)* 2005, **44**:465-468.
6. Van Riel PC, Fransen J: **DAS28: a useful instrument to monitor infliximab treatment in patients with rheumatoid arthritis.** *Arthritis Res Ther* 2005, **7**:189-190.
7. Aletaha D, Ward MM, Machold KP, Nell VP, Stamm T, Smolen JS: **Remission and active disease in rheumatoid arthritis: defining criteria for disease activity states.** *Arthritis Rheum* 2005, **52**: 2625-2636.
8. Makinen H, Kautiainen H, Hannonen P, Sokka T: **Is DAS28 an appropriate tool to assess remission in rheumatoid arthritis?** *Ann Rheum Dis* 2005, **64**:1410-1413.
9. Soubrier M, Zerkak D, Dougados M: **Should we revisit the definition of higher disease activity state in rheumatoid arthritis (RA)? [Abstract.]** *Arthritis Rheum* 2004, **50**:S387.