

RAID: a valid tool to quantify the impact of rheumatoid arthritis. But what impact will it have on the core set for trials?

Maarten Boers

In the context of measurement of health status, "impact" is a relatively new term that encompasses the patient's perspective on the disease.¹ The importance of patient-reported outcomes has traditionally been recognised within the rheumatology community, as evidenced by their presence in the core set of end points for rheumatoid arthritis (RA) clinical trials formulated at the first Outcome Measures in Rheumatology Initiative (OMERACT) conference in 1992.² However, patients and researchers present at the OMERACT 6 conference in 2002 noted that the patient perspective was not adequately covered by the three core set measures: pain, physical function and patient global assessment.³ Domains deemed important but not covered included fatigue, sleep quality and well-being.

More recently, researchers from Bristol have suggested that the impact of a disease or its components is dependent on three interdependent concepts: severity, importance and self-management.⁴ Severity and importance are relatively well-known components that have already been incorporated in so-called "personal preference" tools such as the McMaster-Toronto Arthritis (MACTAR) scale that asks patients to rank their five most important activities and subsequently tracks changes in the ability to perform these activities over time.⁵ Self-management is a new angle brought in by the patients that is really self evident once named, but rarely mentioned before let alone measured in trials. For example, the impact of a disease flare in RA could be expressed in: the total severity (how bad is it?); the importance (are joints that I really need affected?); and the ability to cope (can I still do what I need to do?).

Correspondence to Maarten Boers, Department of Epidemiology and Biostatistics, VU University Medical Center, PK 6Z 185, PO Box 7057, 1007 MB Amsterdam, The Netherlands; ed@vumc.nl

Therefore, it is a happy occasion that two papers in this issue document the validity of a new instrument, Rheumatoid Arthritis Impact of Disease (RAID) Score.^{6,7} The initial development of the instrument was previously reported.⁸ Briefly, European patients were polled to derive a list of the seven most important domains or dimensions of the impact of disease. Over 500 patients from all over Europe were then asked to provide a relative weight for each of the domains.

The paper by Gossec *et al* in this issue of the *Annals* reports on the final development and initial validation of the RAID.⁶ The companion paper by Heiberg and colleagues reports on the performance of the RAID in a Norwegian register of RA patients.⁷ Several other studies are presumably in preparation or in press, judging from the list of abstracts published at the EULAR 2010 conference.

The two papers document the careful and stepwise process taken to develop and validate the instrument. The authors posit that the questions posed by the OMERACT Filter⁹ can be answered positively: Truth: does RAID measure what it intends to measure, in an unbiased way? Discrimination: does it discriminate between situations that are of interest? Feasibility: is the instrument easy to apply and to interpret, at low costs?

Although, for the most part, I can agree with the authors' positive assessment, a closer look reveals some problems that may need to be addressed before we fully embrace this measure.

First, there is a problem with redundancy. As this instrument is specifically targeted for use in clinical trials, the question can be raised what RAID adds to the current core set. The core set was designed to comprise the minimum number of end points that should always be measured in RA trials, and it already includes pain, patient global assessment and physical function, as previously

noted. Thus, the RAID adds fatigue, emotional well-being, sleep and coping. Suppose we were to measure both the core set and RAID that also contains part of the core set: some end points would then be reported twice.

To be included in the core set, a measure should add unique information that is necessary to judge the efficacy of a treatment in a trial. After considering the evidence on existing instruments, OMERACT attendees have previously indicated that fatigue should be measured in all trials,¹⁰ that is, become part of the core set. However, an "official" revision of the core set has not taken place. For the other "new" concepts in the RAID, it is as yet unclear what unique information they bring to that which is already included in the core set. Or rather, as both papers show, such information is probably limited, given the quite strong correlations (0.49–0.77 in Gossec *et al*'s paper and 0.73–0.82 in Heiberg *et al*)^{6,7} with other core set measures as well as legacy instruments such as the SF-36. Already today we face a real problem of losing sight of the wood for the trees, as current trials are sliced and diced to produce separate reports on the core set, radiographs, quality of life, costs, etc followed by yearly extension reports until kingdom come. Thus, routine application of the RAID within trials needs further thought.

Second, the instrument shows good discrimination both in its reliability over repeated observations, detecting different levels of severity, and by detecting changes (or lack of change) in populations where such (lack of) change is expected. Interestingly, simple numerical rating scales performed as well as more extensive instruments in several of the domains. However, I do have a problem with the separate weights suggested for each of the RAID domains. Based on the mean weight given to the seven domains by over 500 patients, the final RAID score has to be calculated by multiplying the score on each of the domains by a separate weighting factor. This increases the face value of the resulting summary RAID score, as it reflects the relative importance of the domains as judged by the patients. But wait... If we look at the 2009 report, the mean weights ranged from 9 to 21 (on a scale of 0–100), with SDs that exceeded that mean in almost every instance.⁸ This indicates a skewed distribution, as borne out by the medians that ranged from 7 to 15; also, five of the seven domains

had the same median weight of 12. This tells us three things: first that the weight given to each of the dimensions varies considerably between patients; second, that some patients give extreme weights to one or two domains, suggesting that the mean may not be the best representation of where the majority of the patients score; and third, that the medians of the weights given to the domains are (very) close together. And even if we were to accept the mean weight as a valid representation of the results, one can question whether the distribution of weights is not simply due to the play of chance. A two-way ANOVA of the 2009 results might give us the answer.

During the peer-review process, the authors were requested to perform analyses of the unweighted RAID sum score. These analyses showed, in some contrast to the statements made in the current report,⁷ that the results of the unweighted score were virtually superimposable upon the results of the weighted RAID score. In all, an unweighted score appears to have comparable face validity, with measurement properties in terms of discrimination that are equal to the weighted RAID score for all intents and purposes.

Quite often, the bottom line for the wide adoption of a measure lies in its

feasibility, the third element of the OMERACT Filter.⁹ The RAID team has made great strides in showing us that each of these important domains can be measured with a simple numerical rating scale. Each of the RAID questions has now been translated into a wealth of languages, respecting cultural context. Therefore, it is a pity that the authors currently still recommend calculating the RAID with this difficult set of weights. However, history has shown that researchers and clinicians are creative in applying an instrument to fit their current needs, and the data suggest that a simple summed RAID score will also serve us well.

In all, we should celebrate the addition of this important instrument to our armamentarium.

Provenance and peer review Commissioned; externally peer reviewed.

Accepted 14 March 2011

REFERENCES

1. **Sanderson T**, Kirwan J. Patient-reported outcomes for arthritis: time to focus on personal life impact measures? *Arthritis Rheum* 2009;**61**:1–3.
2. **Boers M**, Tugwell P, Felson DT, *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 Suppl 1994;**41**:86–9.

3. **Kirwan J**, Heiberg T, Hewlett S, *et al*. Outcomes from the Patient Perspective Workshop at OMERACT 6. *J Rheumatol* 2003;**30**:868–72.
4. **Sanderson TC**, Hewlett SE, Flurey C, *et al*. The impact triad (severity, importance, self-management) as a method of enhancing measurement of personal life impact of rheumatic diseases. *J Rheumatol* 2011;**38**:191–4.
5. **Tugwell P**, Bombardier C, Buchanan WW, *et al*. The MACTAR Patient Preference Disability Questionnaire—an individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. *J Rheumatol* 1987;**14**:446–51.
6. **Gossec L**, Paternotte S, Aanerud GJ, *et al*. Finalisation and validation of the Rheumatoid Arthritis Impact of Disease (RAID) score, a patient-derived composite measure of impact of rheumatoid arthritis. A EULAR initiative. *Ann Rheum Dis* 2011;(In press).
7. **Heiberg T**, Austrad C, Kvien TK, *et al*. Performance of the RAID etc. *Ann Rheum Dis* Published Online First: 6 March 2011. doi:10.1136/ard.2010.143032.
8. **Gossec L**, Dougados M, Rinceval N, *et al*. Elaboration of the preliminary Rheumatoid Arthritis Impact of Disease (RAID) score: a EULAR initiative. *Ann Rheum Dis* 2009;**68**:1680–5.
9. **Boers M**, Brooks P, Strand CV, *et al*. The OMERACT filter for outcome measures in rheumatology. *J Rheumatol* 1998;**25**:198–9.
10. **Kirwan JR**, Minnock P, Adebajo A, *et al*. Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. *J Rheumatol* 2007;**34**:1174–7.

Corrections

Maarten Boers. RAID: a valid tool to quantify the impact of rheumatoid arthritis. But what impact will it have on the core set for trials? *Ann Rheum Dis* 2011;**70**:884–885 doi:10.1136/ard.2010.145375. In the published article, the authors email address is incorrect. The correct email address should be published as: eb@vumc.nl

Ann Rheum Dis 2011;**70**:1349. doi:10.1136/ard.2010.145375corr1