Reappraisal of OMERACT 8 Draft Validation Criteria for a Soluble Biomarker Reflecting Structural Damage Endpoints in Rheumatoid Arthritis, Psoriatic Arthritis, and Spondyloarthritis: The OMERACT 9 v2 Criteria

WALTER P. MAKSYMOWYCH, ROBERT LANDEWÉ, PAUL-PETER TAK, CHRISTOPHER J. RITCHLIN, MIKKEL ØSTERGAARD, PHILIP J. MEASE, HANI EL-GABALAWY, PATRICK GARNERO, DAFNA D. GLADMAN, OLIVER FITZGERALD, DANIEL ALETAHA, VIVIEN P. BYKERK, JOAN M. BATHON, SILJE W. SYVERSEN, MAARTEN BOERS, PIET GEUSENS, ROBERT D. INMAN, VIRGINIA B. KRAUS, TORE K. KVIEN, WILLIAM J. TAYLOR, GEORGE A. WELLS, and DÉSIRÉE van der HEIJDE

ABSTRACT. *Objectives.* A draft set of criteria for the validation of soluble biomarkers reflecting damage endpoints was proposed at OMERACT 8. At OMERACT 9 we aimed to scrutinize the necessity for each of these criteria according to the objectives of the working group.

Methods. The OMERACT 8 draft criteria and the principle objectives of the validation process were clarified at a meeting of the working group in London, December 2007. A new framework was proposed after the following steps were conducted: (A) A systematic review of the literature focusing on the draft criteria and a preselected group of biomarkers (MMP3, CTX-II, RANKL, OPG, CTX-I) followed by a Delphi consensus exercise addressing the importance of individual criteria and identification of omissions in the draft set. (B) Formal debate as well as group discussion centered on the key arguments for inclusion/exclusion of specific criteria. (C) Onsite interactive electronic voting on the importance of specific criteria. The framework was presented and discussed at OMERACT 9 in both breakout and plenary sessions followed by a vote on its acceptance.

Results. The objectives of rheumatoid arthritis, psoriatic arthritis, and ankylosing spondylitis biomarkers in relation to their predictive validity for damage endpoints was clarified and supported by OMERACT 9 participants. The OMERACT 8 draft validation criteria were reformulated into an essential category focused on criteria addressing the OMERACT Filter elements of discrimination (incorporating truth) and feasibility, and a desirable but nonessential category of other criteria addressing truth. This revised draft set was endorsed by participants at OMERACT 9.

Conclusion. A revised set of validation criteria has been drafted by consensus at OMERACT 9 that focuses on the performance characteristics of biomarker assays, the importance of addressing potential confounders, and the essential requirement for clinical validation studies. (J Rheumatol 2009; 36:1785–91; doi:10.3899/jrheum.090346)

Key Indexing Terms:
RHEUMATOID ARTHRITIS   PSORIATIC ARTHRITIS   ANKYLOSING SPONDYLITIS
BIOMARKERS   VALIDATION CRITERIA   STRUCTURAL DAMAGE

From the Department of Medicine, University of Alberta, Edmonton, Alberta, Canada; Division of Clinical Immunology and Rheumatology, Academic Medical Center, University of Amsterdam, The Netherlands; Clinical Immunology Research Center, University of Rochester Medical Center, Rochester, New York, USA; Departments of Rheumatology, Copenhagen University Hospitals at Herlev and Hvidovre, Copenhagen, Denmark; Swedish Medical Center, Seattle, Washington, USA; University of Manitoba, Winnipeg, Manitoba, Canada; INSERM Research Unit 664 and CCBR-SYNARC, Lyon, France; University of Toronto, Toronto Western Research Institute, University Health Network, Toronto, Canada; St. Vincent's University Hospital, Dublin, Ireland; Division of Rheumatology, Department of Internal Medicine III, Medical University of Vienna, Vienna, Austria; University of Toronto, Rebecca MacDonald Center for Arthritis and Autoimmune Diseases, Mount Sinai Hospital, University Health Network, Toronto, Canada; Division of Rheumatology, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA; Department of Rheumatology, Diakonhjemmet Hospital, University of Oslo, Oslo, Norway; Department of Clinical Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands; Internal Medicine/Rheumatology, Maastricht University Medical Centre, The Netherlands and Biomedical Research Centre, University Hasselt, Belgium; University of Toronto, Toronto Western Research Institute, Toronto, Canada; Duke University Medical Center, Durham, North Carolina, USA; Diakonhjemmet Hospital, University of Oslo, Oslo, Norway; University Hospital Maastricht, Maastricht, The Netherlands; Rehabilitation Teaching and Research Unit, University of Otago, Wellington, New Zealand; Department of Epidemiology and Community Medicine, University of Ottawa; Ottawa Health Research Institute, Ottawa, Canada; and Leiden University Medical Center, Leiden, The Netherlands.

Walter P. Maksymowych is a Scientist of the Alberta Heritage Foundation for Medical Research.

W.P. Maksymowych, FRCPC, Professor, Department of Medicine, University of Alberta; P-P. Tak, MD, PhD, Professor of Medicine, Division of Clinical Immunology and Rheumatology, Academic Medical Center, University of Amsterdam; C.T. Ritchlin, MD, PhD, Director, Clinical Immunology Research Center, Professor of Medicine, University

of Rochester Medical Center; M. Østergaard, MD, PhD, DMSc, Professor, Departments of Rheumatology, Copenhagen University Hospitals at Herlev and Hvidovre; P.J. Mease, MD, Professor, Swedish Medical Center; H. El-Gabalawy, MD, FRCPC, Professor of Medicine and Immunology, University of Manitoba; P. Garnero, PhD, DSc, INSERM Research Unit 664 and CCBR-SYNARC; D.D. Gladman, MD, FRCPC, Professor of Medicine, University of Toronto, Senior Scientist, Toronto Western Research Institute, University Health Network; O. FitzGerald, MD, FRCPI, FRCP (UK), Newman Clinical Research Professor, St. Vincents University Hospital; D. Aletaha, MD, PhD, Associate Professor, Division of Rheumatology, Department of Internal Medicine III, Medical University of Vienna; V.P. Bykerk, MD, FRCPC, Assistant Professor of Medicine, University of Toronto, Rebecca MacDonald Center for Arthritis and Autoimmune Diseases, Mount Sinai Hospital, University Health Network; J.M. Bathon, Professor of Medicine, Division of Rheumatology, Department of Medicine, Johns Hopkins University School of Medicine; S.W. Syversen, MD, Department of Rheumatology, Diakonhjemmet Hospital, University of Oslo; M. Boers, MSc, MD, PhD, Professor of Clinical Epidemiology, Department of Clinical Epidemiology and Biostatistics, VU University Medical Center; P. Geusens, MD, PhD, Professor, Internal Medicine/Rheumatology, Maastricht University Medical Centre, Biomedical Research Centre, University Hasselt; R.D. Inman, MD, FRCPC, Professor of Medicine and Immunology, University of Toronto, Senior Scientist, Toronto Western Research Institute; V.B. Kraus, MD, PhD, Associate Professor of Medicine, Duke University Medical Center; T.K. Kvien, MD, PhD, Professor of Rheumatology, Diakonhjemmet Hospital, University of Oslo; R. Landewé, MD, PhD, Professor of Medicine, University Hospital Maastricht; W.J. Taylor, PhD, FRACP, FAFRM, Associate Professor, Rehabilitation Teaching and Research Unit, University of Otago, Wellington; G. Wells, MSc PhD, Professor, Department of Epidemiology and Community Medicine, University of Ottawa, Senior Scientist, Ottawa Health Research Institute; D. van der Heijde, MD, PhD, Professor of Medicine, Leiden University Medical Center.

Address reprint requests to Dr. W.P. Maksymowych, 562 Heritage Medical Research Building, University of Alberta, Edmonton, Alberta T6G 2S2, Canada. E-mail: walter.maksymowych@ualberta.ca

Validation criteria for soluble biomarkers that reflect structural damage endpoints in rheumatoid arthritis (RA), spondyloarthritis (SpA), and psoriatic arthritis (PsA) were drafted by the Soluble Biomarker special interest group (SIG) and evaluated by participants at OMERACT 8[1]. This specific topic constitutes a high priority area for the drug discovery process because structural damage outcomes are important targets for phase III clinical trials. According to the nomenclature adopted by the US National Institutes of Health, structural damage can be considered a candidate surrogate marker predicting longterm (patient-centered) outcome such as disability and death[2]. Development of biomarkers that predict structural damage is attractive, because their availability early in the life cycle of a therapeutic can help decide whether and how to proceed with pivotal clinical trials. From the clinician's perspective, availability of such biomarkers may permit targeting of pre-radiographic disease, and the identification of subgroups at particular risk of disease progression, and may consequently lead to earlier and more aggressive therapeutic intervention in routine clinical practice.

The first draft set of criteria were developed following an electronically conducted Delphi exercise prior to OMER-ACT 8 and then discussed among participants of the SIG at OMERACT 8. The first draft set comprised 14 validation criteria categorized under the OMERACT filter headings of truth, discrimination, and feasibility. Issues highlighted for additional discussion in the further development of the validation criteria included a reappraisal of the relative importance of individual criteria, development of systematic and standardized approaches to the validation of biomarkers at the level of individual criteria, and generation of a levels of evidence template. The OMERACT 8 discussion also identified confusion in the understanding of the primary objectives of the biomarker validation process and in what manner this might differ for the 3 diseases in question. These issues were incorporated into a post-OMERACT 8 agenda, and the group was expanded to an OMERACT working group that included additional investigators active in biomarker and outcomes research. A meeting of the OMERACT biomarker working group was convened in November 2007 in London, England, to clarify the objectives of the working group and formulate proposals that would be addressed further at OMERACT 9.

Our report presents the discussions of the working group. The primary objectives of the validation process are clarified, the development of the proposal for the revised draft set of validation criteria is described, and the deliberations and voting at OMERACT 9 are summarized.

METHODS

As a first step in the London meeting, the objectives of the validation process were clarified. Members were asked to consider the properties of a clinically useful biomarker as it might be used in clinical trials research and in clinical practice to further the understanding of damage progression. Bone mineral density as a biomarker reflecting fracture endpoints was discussed as one such example. A schematic was generated summarizing the properties of a clinically useful biomarker in RA, PsA, and AS. The agreed upon objectives and schematics were then discussed at OMERACT 9.

The OMERACT 8 draft criteria were discussed at the London meeting in both debate and group format followed by an interactive Delphi voting exercise conducted with the KEEpad® system. Arguments for and against inclusion of each of the criteria comprising the OMERACT 8 draft set were presented at the meeting in debate format, with one debater arguing for and one against inclusion of each individual criterion. General discussion ensued. This was followed by the presentation of results of an electronic voting exercise that had 2 objectives: (1) to appraise the importance of individual criteria comprising the 14-criteria draft set proposed for validation of a soluble biomarker as reflecting structural damage endpoints in RA, PsA, AS at OMERACT 8[3]; (2) to identify omissions in the draft criteria as a prelude to drafting additional criteria. The methodology and results are summarized in a companion report[3]. Meeting participants were presented with mean (SD), median (interquartile range), and range scores for the electronic voting exercise. Revisions and additions to the criteria were formulated by consensus.

An interactive Delphi consensus voting exercise on the revised criteria was then conducted onsite using the KEEpad electronic voting system after each individual criterion had been scrutinized in this manner. The question posed to participants was: "Please rate on a scale of 0 (definitely exclude) to 9 (definitely include) to what degree you consider this a required criterion in the validation of a biomarker reflecting structural damage endpoints in: A. Rheumatoid Arthritis; B. Psoriatic Arthritis; C. Ankylosing Spondylitis"

Up to 3 rounds of voting were possible for each disease category with a prespecified mean score of ≤ 3 in any round of voting leading to exclusion and a vote of ≥ 7 in any round of voting leading to inclusion of that specific criterion. The revised criteria set was then presented to OMERACT 9 participants and discussed at breakout group sessions. These discussions and recommendations were summarized at the subsequent report-back plenary session by the rapporteur. All participants were then asked to vote on the following question: "The working group has proposed OMERACT 9 v2 criteria based on a core set and a desirable but not essential category. Do you agree with these criteria and this framework?" Consensus vote in favor of the proposition was predefined as ≥ 70% of voters agreeing with the proposition.

RESULTS AND DISCUSSION

Objectives of the validation process. The following primary objectives for an RA biomarker were proposed by the biomarker working group in London and discussed at OMERACT 9.

A. *Change* in the biomarker reflects/predicts *change* in the damage endpoint at the level of:

Group: The biomarker constitutes an endpoint for clinical trials and cohort studies.

Individual patient: The biomarker constitutes an endpoint for clinical practice.

B. *Change* in the biomarker reflects/predicts *change* in the damage endpoint *independently* of known predictors, e.g., Disease Activity Score, baseline damage, rheumatoid factor, anti-cyclic citrullinated peptide, shared epitope, C-reactive protein (CRP)/erythrocyte sedimentation rate.

The primary objectives for a PsA biomarker were proposed to be the same as for RA. The following primary objective was proposed by the group for an AS biomarker:

Change in biomarker predicts *change* in damage endpoint independently of known predictors, e.g., baseline damage.

No revisions to these objectives were proposed at the OMERACT 9 meeting. In retrospect it was recognized that the OMERACT Filter criteria of truth and discrimination largely overlap in the setting of validation of predictive biomarkers; this is in contrast to the setting of validation of evaluation measures, where the Filter was originally developed[4]. The general question attached to the element of Truth is: "does it measure what it's supposed to?" and in the setting of prediction this translates to: "Does it adequately predict what it should predict?" Formulated in this way, the question is very similar to that posed for the element of Discrimination: "Does the measure distinguish between states of interest?" If the biomarker is modifiable (e.g., a soluble biomarker) and change in this biomarker over the course of a longitudinal study or clinical trial is shown to consistently parallel change in radiographic progression as disease activity changes and/or with changes in therapy, it can be stated that the biomarker reflects damage to the extent that it no longer becomes necessary to measure radiographic progression. To be clinically useful, such a biomarker should reflect radiographic progression in groups of patients assessed in longitudinal studies and clinical trials as

well as in the individual patient, thereby allowing the clinician to manage therapy according to prognostic risk.

Properties of a clinically useful RA or PsA biomarker (Figure 1). Change in the damage endpoint can occur relatively rapidly in RA and has been documented as soon as 3 months in patients in a placebo section of a trial[4]. To be clinically useful, an RA or PsA biomarker should be a leading indicator of change in the damage endpoint, and the change in biomarker should correlate with the interval change in damage progression. Moreover, the magnitude of the change in biomarker level should consistently reflect the subsequent degree of change in radiographic progression, whether this is associated with spontaneous or treatment-induced change in disease activity. Since the biomarker should reflect radiographic progression independently of known predictors, such as the Disease Activity Score (DAS), changes in biomarker levels may not necessarily correlate with changes in other predictors, and the magnitude of change may vary in relation to change in other predictors. However, from the perspective of clinical utility it would be desirable if such a biomarker was more responsive than routinely assessed clinical and laboratory measures associated with radiographic progression. Finally, measurement of the biomarker should add prognostic information regarding radiographic progression over and above the combined information obtained from all other known predictors. These properties should not only be demonstrable at the group level, as in the evaluation of response to treatment in clinical trials and cohort studies, but also at the individual patient level. Figure 1 illustrates these principles in this hypothetical example of a clinically useful RA biomarker. At baseline, biomarker level, DAS, and radiographic progression are high. Treatment with methotrexate is associated with an immediate reduction in biomarker level in the first 3 months that is much greater than the reduction in DAS and precedes any change in radiographic progression that decelerates from 3–6 months. From 6 months there is a very large increase in biomarker level that precedes (i.e., predicts) a large increase in radiographic progression observed after 9 months. At the one-year point the anti-tumor necrosis factor (TNF) treatment is added to methotrexate and this is associated with an immediate and very large decrease in biomarker level from one year to 15 months that precedes the large reduction in radiographic progression observed from 15 months. Note that the magnitude of the change in the RA biomarker level reflects the subsequent change in the degree of radiographic progression and that the biomarker is also much more responsive than the DAS.

Properties of a clinically useful AS biomarker (Figure 2). Validation of a biomarker for radiographic progression in AS presents different challenges because radiographic progression is slow, requiring at least 2 years of followup before progression can be reliably measured, because there is as yet no demonstrable association between disease activ-
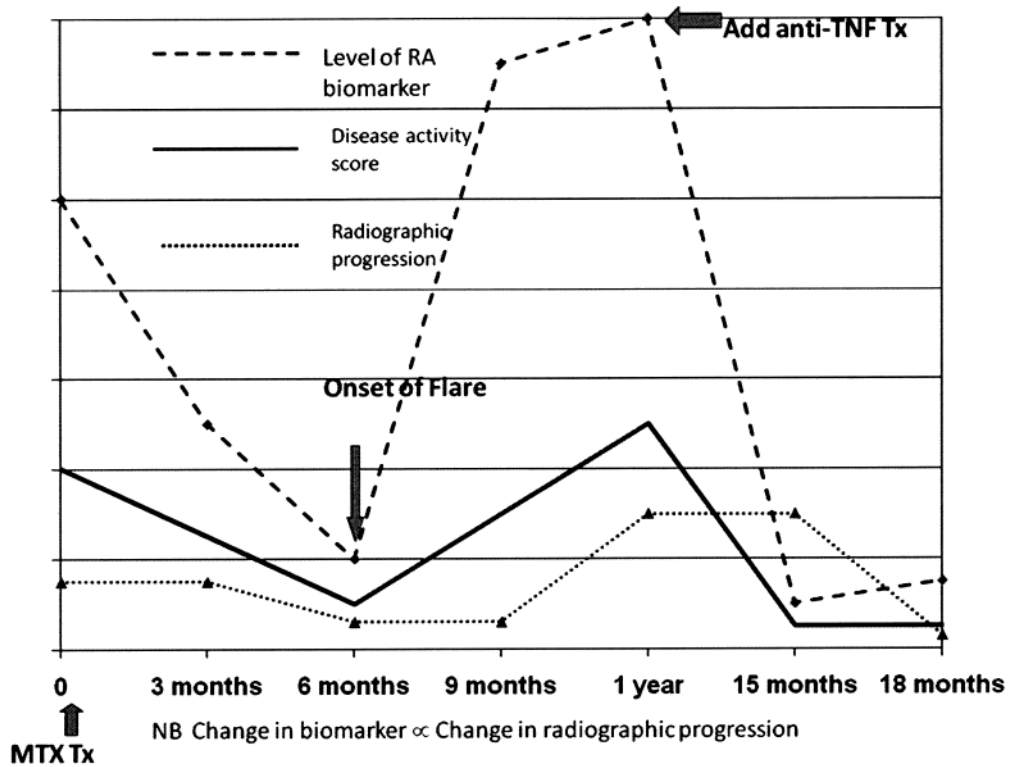
Figure 1. The properties of an "ideal" RA biomarker that can substitute for radiographic damage in clinical studies and individual patients with RA. The relationship between levels of the biomarker, Disease Activity Score (DAS), and the interval change in radiographic severity over 3 months (radiographic progression) is illustrated. Tx: therapy.
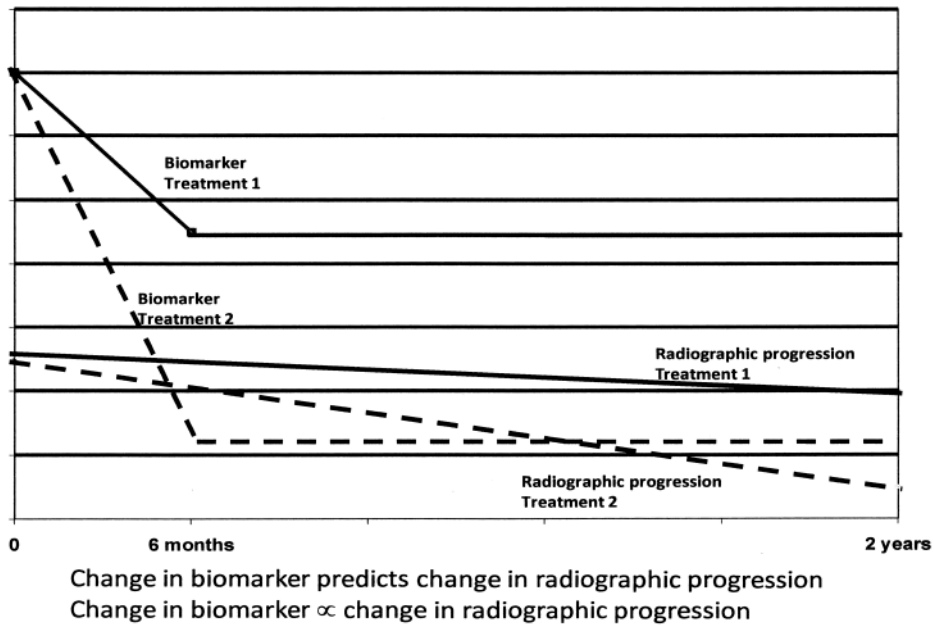

Change in biomarker predicts change in radiographic progression
Change in biomarker ∝ change in radiographic progression

Figure 2. How a "response-to-treatment" effect of an AS biomarker can provide leading indicator prognostic data for radiographic damage in AS. Treatment 2 reduces biomarker levels at 6 months to a greater degree than treatment 1, which predicts less radiographic progression over 2 years with treatment 2.

ity measures and radiographic progression, and because there is currently no good evidence that any treatment slows progression in this disease[5,6]. An AS biomarker that predicts damage progression may, therefore, vary unpredictably over the long time-course of observation. Analyses that address this variability over the duration of followup, for example area under the curve, may still define an important association between the biomarker and damage progression in cohort studies. However, the conclusions from this type of analysis may not be useful to the clinician managing the individual patient. Assessment of the predictive validity of a change in biomarker level early in the course of followup is clinically meaningful, although this may be problematic when the factors that govern progression are largely unknown, and biomarker levels may fluctuate unpredictably. This challenge might be addressed in a "response to treatment" study design using an intervention that has been shown to induce a consistent and persistent change in the biomarker, e.g., reduction in matrix metalloprotease-3 induced by anti-TNF agents[7]. The magnitude of such a change should reflect the subsequent change in risk for radiographic progression and could be examined at both the group and the individual patient level. This approach would require preliminary studies addressing the influence of treatment on the biomarker. Figure 2 illustrates these principles in this example of a clinically useful AS biomarker in 2 groups of patients that have the same radiographic progression at baseline but are exposed to 2 different treatments, with treatment 2 having a greater effect on the AS biomarker than treatment 1. The greater reduction in the level of the AS biomarker with treatment 2 predicts a greater reduction in radiographic progression.

Reappraisal of OMERACT 8 draft criteria. The following conclusions were reached by the group at the London meeting in developing a revised criteria set.

1. Voting for inclusion/exclusion of individual criteria was decisive in that consensus was achieved after the first round of voting for all the criteria. Nevertheless, even those criteria that were voted for exclusion were still considered desirable in developing a body of scientific evidence in support of a biomarker as reflecting joint damage. Consequently, the group proposed that the criteria be divided into 2 categories, an essential or core set and a desirable but nonessential category (see Appendix).

2. The 5 criteria organized under the category of truth were considered desirable but not essential criteria. The original wording of Criterion 2 was revised to remove the word "immunohistochemically" in describing localization to joint tissues. The group felt that there are biomarkers that may be useful but have not been specifically localized to joint tissue using immunohistochemistry.

Conclusions are often drawn too early about the biology and mechanism of action of biomarkers from work conducted in animal models. This can lead to misinterpretation of the data in humans. Animal models may be useful for under-standing pathophysiology but are not useful as a validation tool. A biomarker may be localized to joint tissues by several different methodologies of which immunohistochemistry is one example. But few demonstrate specificity for target of joint tissue origin and some clinically useful biomarkers are not primarily of joint origin, e.g., CRP. Requiring an understanding of the relation between a biomarker and turnover of joint tissue components is problematic, because there is no recognized *in vivo* "gold standard" for bone/cartilage turnover and *in vitro/ex vivo* models of bone/cartilage turnover are poorly predictive of *in vivo* relationships. *In vivo* techniques are also invasive and therefore not feasible. With respect to demonstrating correlations with other surrogates reflecting damage endpoints this criterion might be appropriate if magnetic resonance imaging is used, because increasing data supports its predictive validity for radiographic damage endpoints. It would not be appropriate for other biomarkers such as CRP because of weak association with damage.

3. Under the category of discrimination, Criterion 6 was revised as follows: "The assay for measurement of the biomarker is reproducible (coefficient of variation: intra-assay less than 10%, inter-assay less than 15%)" was changed to "inter-assay less than 10%," as it was felt to be a more stringent cutoff.

4. Criterion 7 describing sources of variability on levels of the biomarker was revised to stipulate core variables that should always be examined during the validation process (age, sex, ethnicity, circadian rhythms, body mass index, renal/hepatic function, fasting/nonfasting) and desirable but nonessential variables (menopause, comorbidity, physical activity, nonsteroidal antiinflammatory drugs). This criterion was considered important but lacked feasibility in its original form because of its complexity in requiring examination of so many variables.

5. Criteria 8 and 9 under the category of discrimination referring to metabolism of the biomarker and sensitivity/specificity of the biomarker in the disease population compared to healthy controls, respectively, were considered desirable but nonessential.

Metabolism, clearance, and half-life of a biomarker are not relevant with respect to its predictive validity because what is being tested is the strength of the association between change in levels of a biomarker and change in damage endpoints. Its pharmacokinetic behavior in healthy and affected individuals is, therefore, not relevant to its validation for this objective. Specificity is also not relevant if the biomarker is being used as a predictor of outcome in affected individuals and is of greater importance if the purpose of the biomarker is diagnostic.

6. Criteria 10, 11, and 12 require that the biomarker demonstrates independent association with the radiographic endpoint in patients in (i) prospective cohort studies (Criterion 10); (ii) randomized controlled trials (Criterion 11); and (iii)

pre-radiographic disease (Criterion 12). The 3 criteria were combined to state that the association be evident "at all stages of disease at the level of both absolute and relative change in prospective cohort studies and randomized controlled trials (except for AS) of adequate sample size, and followed for a sufficient duration to detect change in x-ray damage score." This was also modified in recognition that validation of an AS biomarker in a randomized controlled trial would not be feasible using a plain radiographic endpoint in view of the necessity to conduct a 2-year trial.

7. Criterion 13, under feasibility, was modified to delete the requirement for methodological simplicity and to emphasize accessibility of the biomarker assay. International standardization of an assay was thought to be a difficult but desirable criterion. The statement was further modified to remove the criterion for the assay to be "well characterized" as this was considered vague.

8. Criterion 14 was modified to address the requirement for stability after repeat freeze/thaw cycles and after longterm storage (> 1 year).

Several additional revisions to the criteria were proposed at the OMERACT 9 meeting. For Criterion 6, addressing assay reproducibility, it was recommended that this be conducted through formal reliability analysis. For Criterion 7, describing sources of variability, the importance of using appropriate controls as compared to merely healthy individuals was proposed to address demographic confounders. For Criterion 13, addressing accessibility of the biomarker

---

Appendix. OMERACT 9 draft validation criteria for a soluble biomarker to be regarded as a valid biomarker reflecting structural damage in rheumatoid arthritis (RA),psoriatic arthritisi (PsA), and spondyloarthritis.

Essential Criteria
Truth and Discrimination
The assay for measurement of the biomarker is reproducible according to reliability analysis.
The effects of the following sources of variability on levels of the biomarker in appropriate controls are known:

Core Variables	**Desirable but Not Essential Variables**
Age	Menopause
Sex	Physical activity
Ethnicity	Nonsteroidal antiinflammatory drugs
Circadian rhythms	Comorbidity
Body mass index	
Renal/hepatic function	
Fasting/non-fasting	

The biomarker demonstrates independent association with the structural damage endpoint (van der Heijde modification of Sharp score for RA/PsA, modified Stoke AS Spinal Score for AS) in patients at all stages of disease at the level of both absolute and relative change in prospective cohort studies and randomized controlled trials* of adequate sample size, and followed for a sufficient duration to detect change in radiographic damage score.
Feasibility
The assay for measurement of the biomarker is internationally standardized (availability of reference standards), and is readily accessible if used for clinical practice.
Stability of the biomarker at room temperature, in frozen specimen, after repeat freeze/thaw cycles, and after longterm storage has been documented.
Desirable but Non-essential Criteria
Truth
A preclinical body of evidence that the biomarker reflects tissue remodelling in established animal models of disease (e.g., collagen arthritis for RA).
• The biomarker has been localized to joint tissues.
• The biomarker demonstrates sensitivity and specificity for target of joint tissue origin.
• Relation of biomarker to synthesis, degradation, turnover of joint tissue components has been characterized.
• Levels of the biomarker correlate with scores for other surrogates that have been established as possessing predictive validity for structural damage (e.g., MRI for erosive RA).
Discrimination
The metabolism, clearance, and half-life of the biomarker have been characterized in normal individuals and in patients with arthritis. The impact of concomitant medications has been examined.
The biomarker demonstrates high sensitivity and specificity in comparisons of the disease population with age and sex matched healthy controls.

---

*Not feasible for AS.

assay, the particular relevance of this to clinical practice was stated. A single revision was proposed to one of the nonessential criteria addressing the metabolism of the biomarker (Criterion 8), where the stipulation was added that the impact of concomitant therapy should be determined. The percentage of OMERACT 9 participants attending the workshop who voted "yes" on the question of agreement with this framework and the specific criteria was 79%.

CONCLUSIONS

The revised criteria focus squarely on the feasibility and discrimination components of the OMERACT filter in the approach to validation studies aimed at biomarkers reflecting damage endpoints. Particular attention needs to focus on the performance characteristics of biomarker assays, the importance of addressing potential confounders, and the essential requirement for clinical validation studies. There has been little discussion addressing the principal requirements of prognostic studies in arthritis and we hope that these criteria will stimulate further discussion aimed at a minimum set of requirements for study design, as has been developed for therapeutic interventions. Moreover, it is our hope that manufacturers of biomarker assays will heed the essential requirements for biomarker performance stipulated in these criteria and will provide the scientific community with the information that is so often lacking in the supplied assay literature.

REFERENCES

1. Maksymowych WP, Landewé R, Poole AR, et al. Development of draft validation criteria for a soluble biomarker to be regarded as a valid biomarker reflecting structural damage endpoints in rheumatoid arthritis and spondyloarthritis clinical trials. J Rheumatol 2007;34:634-40.
2. De Gruttola VG, Clax P, DeMets DL, et al. Considerations in the evaluation of surrogate endpoints in clinical trials. Summary of a National Institutes of Health workshop. Control Clin Trials 2001;22:485-502.
3. Syversen SW, Landewé R, van der Heijde D, et al. Testing of the OMERACT 8 draft validation criteria for a soluble biomarker reflecting structural damage in rheumatoid arthritis: A systematic literature search on 5 candidate biomarkers. J Rheumatol 2009;36:1769-84.
4. Bruynesteyn K, Landewé R, van der Linden S, van der Heijde D. Radiography as primary outcome in rheumatoid arthritis: acceptable sample sizes for trials with 3 months follow-up. Ann Rheum Dis 2004;63:1413-8.
5. Wanders AJ, Landewé RB, Spoorenberg A, et al. What is the most appropriate radiologic scoring method for ankylosing spondylitis? A comparison of the available methods based on the Outcome Measures in Rheumatology Clinical Trials filter. Arthritis Rheum 2004;50:2622-32.
6. van der Heijde DM, Wanders A, Mielants H, Dougados M, Landewé RBM. Prediction of progression of radiographic damage over 4 years in patients with ankylosing spondylitis [abstract]. Ann Rheum Dis 2004;63:OP132.
7. Maksymowych WP, Rahman P, Shojania K, et al. Beneficial effects of adalimumab on biomarkers reflecting structural damage in patients with ankylosing spondylitis. J Rheumatol 2008;35:2030-7.