

Economic Evaluation of Programs or Interventions in the Management of Rheumatoid Arthritis: Defining a Consensus-based Reference Case

ANDREAS MAETZEL, PETER TUGWELL, MAARTEN BOERS, FRANCIS GUILLEMIN, DOUG COYLE, MIKE DRUMMOND, JOHN B. WONG, and SHERINE E. GABRIEL, on behalf of the OMERACT 6 Economics Research Group

ABSTRACT. Improvement in the quality of economic evaluation could be documented as a consequence of international and national standardization efforts. One such effort is the recommendation that all economic evaluations in a given field produce findings in a standard format using a reference case. A reference case-based economic evaluation would adhere to specific settings with regard to outcomes, comparators, modeling techniques, and use of costs to facilitate comparisons among economic evaluations performed with the same objective. In the past, the Outcome Measures in Rheumatology Clinical Trials (OMERACT) consensus conference has successfully developed widely used, consensus-based outcome criteria for clinical improvement in rheumatoid arthritis (RA). Present efforts are being directed at the development of recommendations for the type and format of a reference case economic evaluation for newly developed disease modifying antirheumatic drugs (DMARD). This document discusses 13 important elements that experts considered to be relevant for the development of a reference case recommendation for economic evaluations in RA. We provide the rationale for each element and discuss how each element has been addressed in published economic evaluations of DMARD. (J Rheumatol 2003;30:891–6)

Key Indexing Terms:

RHEUMATOID ARTHRITIS

ECONOMIC EVALUATION

STANDARDS

Cost-effectiveness analyses assist policy makers and budget holders to better understand the economic consequences of reimbursement decisions for new therapies or technologies. However, cost-effectiveness analyses do not and cannot form the sole basis for such decisions. This is due, in part, to

the well documented equity concerns and to potential methodological shortcomings of current economic analyses^{1,2}. Fortunately, efforts are under way to strengthen and streamline cost-effectiveness analyses through development of methodological standards³. Making cost-effectiveness analyses adhere to a set of basic methodological standards not only ensures a standard of methodological rigor but also enables cost-effectiveness results to be compared across disciplines and disorders. Several sets of standards have been published⁴⁻⁹. The leading standard is the reference case recommendations made by the US Panel on Cost Effectiveness in Health and Medicine⁵⁻⁷. The scientific community has largely accepted these recommendations. Indeed a recent appraisal of published cost-effectiveness analyses based its assessment of quality on whether studies were “panel-worthy” or not³. If these and other methodological standards are generally adopted, future cost-effectiveness analyses may have a more influential role in policy making regarding new therapies and technologies.

One limitation of current methodological standards is that they are, of necessity, fairly general and tend to neglect the unique circumstances that surround particular clinical contexts. Even with close adaptation of an analysis to existing sets of standards, the design or execution of an analysis could still result in incomplete or even potentially biased comparisons across programs or interventions within certain clinical disorders. Recognizing this problem, and the

From the Division of Clinical Decision Making and Health Care Research, Toronto General Research Institute, Toronto; Clinical Epidemiology Unit, Loeb Research Institute, Ottawa Hospital, Ottawa; Centre for Global Health, University of Ottawa, Institute of Population Health, Ottawa, Canada; Department of Clinical Epidemiology and Biostatistics, VU University Hospital, Amsterdam, The Netherlands; School of Public Health, Faculty of Medicine, Vandoeuvre-les-Nancy, France; Centre for Health Economics, University of York, York, UK; Division of Clinical Decision Making, Informatics and Telemedicine, New England Medical Center, Boston, MA; Department of Health Sciences Research and Division of Rheumatology, Department of Internal Medicine, Mayo Clinic, Rochester, MN, USA.

A. Maetzel, MD, PhD, Division of Clinical Decision Making and Health Care Research, Toronto General Research Institute; P. Tugwell, MD, Clinical Epidemiology Unit, Loeb Research Institute, and Centre for Global Health, University of Ottawa, Institute of Population Health; M. Boers, MD, PhD, Department of Clinical Epidemiology and Biostatistics, VU University Hospital; F. Guillemin, MD, PhD, School of Public Health, Faculty of Medicine; D. Coyle, MSc, Clinical Epidemiology Unit, Loeb Research Institute; M. Drummond, PhD, Centre for Health Economics, University of York; J.B. Wong, MD, Division of Clinical Decision Making, Informatics and Telemedicine, New England Medical Center; S.E. Gabriel, MD, MSc, Department of Health Sciences Research and Division of Rheumatology, Department of Internal Medicine, Mayo Clinic.

Address reprint requests to Dr. A. Maetzel, Division of Clinical Decision Making, University Health Network, 200 Elizabeth Street, EN-232A, Toronto, Ontario M5G 2C4. E-mail: maetzel@uhnres.utoronto.ca

unique challenges surrounding economic analysis of rheumatoid arthritis (RA), the Outcome Measures in Rheumatology Clinical Trials (OMERACT) Organizing Committee convened experts with methodological and clinical expertise to draft guidance documents to help standardize economic evaluations in rheumatological disorders, thereby making them more easily compared through the application of a common structural approach for a particular clinical context¹⁰. These recommendations are intended to supplement the generally accepted methodological standards to assist their application in rheumatology.

This paper focuses on the choices of (1) outcome measures, (2) comparators, (3) modeling techniques, and (4) costs in RA. Within these main categories, the 13 previously identified elements that require further clarification and consensus will be discussed¹⁰. For each identified element we will describe the alternative methodologies, and the degree to which it could possibly be addressed in cost-effectiveness analyses of programs or interventions in RA, and the rationale for the final recommendation. We then discuss these elements in the context of 3 recent examples of cost-effectiveness analyses in RA, i.e., a comparison of stepped-down combination therapy in RA¹¹, a 6-month model-based economic evaluation of disease modifying antirheumatic drugs (DMARD) and biologics in methotrexate (MTX)-resistant RA¹², and an evaluation of the addition of leflunomide to a conventional strategy of DMARD¹³. All 3 economic evaluations are briefly summarized below.

Economic evaluation of stepped-down combination therapy in RA

This evaluation was performed alongside a 56-week randomized, double blind controlled clinical trial of combination therapy with sulfasalazine, MTX, and prednisolone compared to sulfasalazine alone in 156 patients with early, active RA¹⁴. Direct medical and non-medical costs were collected in weekly diaries. The Disease Activity Score (DAS) was measured and utilities were assessed as per the rating scale and standard gamble method. At 56 weeks, combined therapy was less expensive, albeit not significantly, and more efficacious than therapy with sulfasalazine alone¹¹.

Evaluation of DMARD and/or tumor necrosis factor- α antagonists in MTX resistant RA

This was a US-based, 6-month evaluation that compared the tumor necrosis factor- α (TNF- α) antagonist etanercept, either in combination with MTX or alone, to triple therapy (MTX combined with hydroxychloroquine and sulfasalazine), combination of MTX and cyclosporine, and continuation of MTX alone despite failure¹². The incremental cost per patient achieving American College of Rheumatology (ACR) core set 20% improvement was calculated, as well as incremental cost per weighted ACR

20%, 50%, and 70% improvement. Triple therapy was found to be more cost-effective than etanercept in combination with MTX.

Economic evaluation of adding leflunomide to a conventional strategy of DMARD

This evaluation was done from a Canadian public payer's perspective¹³. It was based on a systematic literature review of withdrawal rates of conventional DMARD¹⁵ and supported by surveys of Canadian rheumatologists. It evaluated the cost-effectiveness of adding leflunomide within a 5-year time horizon to a conventional strategy of DMARD, i.e., MTX followed by combination with sulfasalazine and hydroxychloroquine, followed by injectable gold salts and low dose cyclosporine. The evaluation showed that leflunomide would cost about US\$14,000 per year of ACR 20 response gained and US\$72,000 per standard gamble quality adjusted life-years (QALY) gained over a 5-year time horizon.

OUTCOME MEASURES

Element 1: Which clinical outcome measures should we use?

General guidelines for the conduct and reporting of economic evaluation recommend describing benefits in QALY, so that morbidity and mortality consequences are expressed as a single preference-based measure¹⁶. However, participants of the OMERACT 5 conference in Toulouse agreed that disease-related clinical outcomes should also be included in economic evaluations and that both intermediate and final outcomes should be considered¹⁰. Intermediate outcomes, however, should only be included if they demonstrate a strong and consistent relationship to final outcomes.

The ACR recommends the use of relative improvement criteria as outcomes in clinical trials, i.e., the classification of treatment responders according to at least a 20% improvement in a composite measure of clinical, functional, and pain indices (ACR 20 responder)¹⁷. These criteria have now been accepted in almost all recently conducted trials. Both ACR 20 and EULAR criteria for improvement can be applied in economic evaluations, given that they result in similar classifications for trial patients. The EULAR criteria, based on the DAS, require both an absolute improvement in the DAS and the attainment of a DAS level that is associated with low disease activity¹⁸.

Despite the prevailing use of relative response rates, these measures are less suitable as outcomes for longterm modeling of disease (this element is further discussed below). For example, response relative to baseline may have less significance as duration of a particular therapy increases. Describing the disease status in absolute terms may be more relevant in this circumstance. This can be done by a continuous absolute measure such as the DAS. It can also be done by counting the proportion of patients classi-

fied as being in remission or in a low disease activity state. The OMERACT Minimal Clinically Important Difference Task Force is currently developing criteria to define such a low disease activity state. The Health Assessment Questionnaire (HAQ) Disability Index has been used to characterize longterm patient outcomes¹⁹. Incorporating outcome measures that not only describe short term success but that also characterize the natural history of the disease (e.g., the development of disability) over the long term is critically important in economic evaluations in RA. This element is seldom considered in the more generic methodological standards.

How has this been addressed practically. For each possible treatment regimen, Choi and colleagues defined treatment-specific response as a proportion of the maximum achievable response. They assessed the additional or incremental benefit of each treatment by comparing it to the next least expensive but effective treatment¹². This relative measure of response was useful in the short 6-month time horizon of the evaluation¹², but neglects potential longterm benefits and risks. Maetzel and colleagues modeled ACR 20 response as the fraction of patients who continue therapy¹³. The authors hypothesized that while the number of patients maintaining therapy continually declines, the fraction of responders among those who continue stays constant. Of course, while clinically reasonable, this assumption remains hypothetical and requires empirical validation.

Element 2: Which sources for QALY?

Many preference-based measures are available and appropriate. Even indirect measures such as the EuroQol-5D²⁰ or the Health Utilities Index Mark III (HUI)²¹ are acceptable and are especially useful for policy makers. It is unclear whether utilities should be derived from the public, as is the case with indirect measures, or directly from the patient, which clinicians seem to prefer.

How has this been addressed practically. Both Verhoeven, *et al*¹¹ and Maetzel, *et al*¹⁵ measured patient preferences with the rating scale and standard gamble methods. In both studies the rating scale assessments were able to significantly differentiate the clinically superior therapy from the comparator, but not so the standard gamble method. These results suggest that standard gamble may be a less sensitive method for eliciting patient preferences.

Element 3: How should adverse events be reported and classified?

Although modeling adverse events is as important as modeling benefits, based on information from both clinical trials and observational studies, there is little consensus on how that should be done. In fact, the reporting and classification of adverse events in both clinical trials and observational studies in RA lack standardization. A classification system of adverse events based on common toxicity criteria

has been proposed by Woodworth, *et al* for rheumatology, but whether this classification adequately captures the cost and the impact on quality of life of adverse events remains unknown²². The ongoing work of the OMERACT Toxicity Group may well address this element.

How has this been addressed practically. The study by Verhoeven, *et al*¹¹ adequately captured the cost and quality of life consequences of adverse events by prospectively tracking such events. Choi and associates referred to published estimates of major and minor toxicity associated with MTX¹², but in the original study, these were derived not from prospective evaluations but from an informal assessment of published observational studies. Similarly, the Maetzel, *et al* metaanalysis of observational studies pooled published incidences of specific clinical adverse events and estimated their associated costs using surveys conducted with experts¹⁵.

Element 4: How should we model mortality that follows major events?

While most clinical trials contain inadequate power or followup to differentiate mortality benefits in RA associated with one intervention versus another, the consensus of methodological experts and clinicians is that mortality data should be included if available.

How has this been addressed practically. None of the studies mentioned above has included mortality as a treatment-specific consequence.

COMPARATORS

The US Panel on Cost Effectiveness in Health and Medicine recommends that the health intervention of interest should be compared to existing practice, i.e., the “standard of care.” If existing practice is not a cost-effective option, then other options should be considered, such as the best available alternative, a viable, low cost alternative, or a “do-nothing” alternative¹⁶. Other publications make similar recommendations. However, it is recognized, particularly by the guidelines of the UK National Institutes of Clinical Excellence, that “absence of head-to-head comparisons is a common problem especially for new drugs when registration trials have been placebo-controlled”⁸.

Element 5: How should we compare drugs in the absence of head-to-head trials?

The absence of head-to-head comparisons is of particular importance in RA, as the characteristics of the patients recruited into clinical trials of DMARD have changed rather dramatically over time. Initially, DMARD treatment was reserved for more severely affected patients in the later stages of the disease. Currently, some clinicians advocate aggressive treatment for all patients with RA, even those early in their disease course, in order to prevent longterm disability. Further, some new drug trials specifically target

patients refractory to at least 2 or 3 DMARD, whereas others recruit only DMARD-naïve patients. Thus, because the patient populations in these new drug trials differ substantially, and because there are no head-to-head trials, making unbiased estimates of the relative benefit of different therapeutic agents compared to each other is not possible. Therefore, in the absence of head-to-head comparisons, any statement regarding the superiority of any drug over another should be viewed as speculative.

How has this been addressed practically. Verhoeven, *et al*¹¹ only compared the drug regimens involved in the trial and did not attempt any indirect comparison of DMARD. Choi, *et al*¹² calculated therapy-specific ACR 20 response as a fraction of the response achievable beyond that of placebo for MTX-resistant patients continuing MTX and thus obtained an estimate of “relative response beyond placebo” for each comparator. However, although data are available for the degree of response among MTX-resistant patients who withdraw from MTX entirely, the degree of response for continued MTX in MTX-resistant patients was based on response observed in one etanercept trial. This single treatment arm formed the basis for comparison for all of the various DMARD strategies. Because of differences in patient populations, the validity of some of the indirect comparisons is subject to debate. On the other hand, their analysis did attempt to address the real-world treatment decisions facing practicing rheumatologists. In the study by Maetzel, *et al*¹⁵, the degree of response was presented as a fraction of those continuing and was thus supported by data from multiple observational studies for almost all comparators (except leflunomide). Clearly, further research into the methodologies to infer benefit from indirect comparisons needs to be performed.

Element 6: Should we model treatment sequences, and if so, how?

The treatment of RA involves not one therapeutic agent but rather a sequence of therapies over the long term, i.e., a therapeutic sequence strategy. Thus, modeling of therapeutic sequences is of particular importance in the economic evaluation of DMARD over the long term. Realistic sequences should be modeled, depending on the clinical setting, but research documenting the type of sequences used by rheumatologists is lacking. New drugs should be evaluated in patient populations similar to those in which clinical trials occurred and thus in the appropriate realistic treatment sequence.

How has this been addressed practically. For example, Maetzel, *et al*¹⁵ modeled the addition of leflunomide within a realistic sequence of DMARD treatment using a survey of Canadian rheumatologists. The relevant realistic treatment sequence may change as new therapies that are now available to treat RA are integrated into standard care. Evaluating the cost-effectiveness of a therapy within a therapeutic

sequence may identify the most clinically relevant population for a new drug and where it may be best positioned within a therapeutic sequence.

MODELING TECHNIQUES

The ability of pharmacoeconomic models to project health outcomes and resource utilization beyond the duration of the underlying clinical studies allows policy makers to examine the relative merits of alternative allocation schemes using the same longterm or lifetime time horizon and thus is an attractive element of economic evaluations. For example, one could examine the potential future benefit of a life-saving intervention, by “crediting” the intervention with the potential future life-years saved instead of limiting clinical benefit to the duration of the clinical trial. Although this may be particularly relevant for RA therapies that prevent future disability, accurately forecasting gains in typical rheumatological outcomes, such as functional status, quality of life, or subjective treatment response, may be problematic. Most general guidelines support the necessity to model beyond what is known *de facto*, but call for caution in the interpretation of those results and demand detailed sensitivity analysis over the range of all underlying assumptions.

The 2 elements that are probably most contentious are, however, of considerable relevance to the modeling of interventions and programs intended for RA: (1) Should we model beyond the duration of clinical trials and if so, how? and (2) Should we model beyond duration of pharmacological therapy when evaluating DMARD? These questions stand out because clinical trials in RA seldom continue longer than one year; yet understanding the course of the disease over the long term is important, in particular with regard to therapy failures, disease remission, progression of structure damage, and delayed side effects.

Elements 7, 8, 9: What should the horizon be? Should therapy be continuous? How should we model beyond trial duration?

Most agree that models should include at least one-year time horizons and that RA therapy should be continuous. However, modeling beyond the duration of the trial that supports the intervention is more contentious. Such models require data on treatment withdrawal and degree of response. Information on the extent of treatment withdrawals, i.e., the fraction of patients discontinuing therapy due to lack of efficacy or adverse events, could be obtained from the supporting clinical trials and strengthened by data from observational studies. However, modeling clinical benefit beyond trial duration is less straightforward because of the need for relative outcome measures. The choice of the clinical outcome measure will affect longterm modeling efforts. One compromise approach that errs on the side of caution would be to use clinical trial data over a time span of one year (the duration of most modern clinical trials in

RA), and a time horizon of 5 or 10 years, since estimates beyond these time horizons may be too uncertain to be clinically acceptable.

How has this been addressed practically. Modeling beyond the trial duration was considered neither in the one-year trial-based study by Verhoeven, *et al*¹¹, nor in the cost-effectiveness analysis by Choi, *et al*¹². The inclusion of a longterm perspective in the study by Verhoeven, *et al*, however, would permit explicit incorporation of future effects of steroids and of halting the progression of structural damage early in the disease.

The evaluation by Maetzel and associates was based on findings from a metaanalysis of withdrawal data from published observational studies and randomized controlled trials of DMARD¹⁵. Response was modeled as a fraction of those who continue, which provided an additional qualification of therapeutic success. Response was then mapped to measures of patient preferences obtained from a randomized controlled trial that compared leflunomide to MTX or placebo. Although the metaanalysis of DMARD withdrawal rates permitted the modeling of DMARD strategies, pooling published data did not permit estimating the absolute decline in health status over time associated with each treatment.

Element 10 and 11: Discontinuation of therapy; extrapolation beyond therapy

Modeling beyond pharmacological therapy is a problematic element that could potentially impede the comparability of economic evaluations. Little is known about how patients who are refractory to all pharmacological interventions fare in the long term. Is there a sudden and steep increase in disease progression, or a major shift to joint-preserving surgery? Does management become very expensive once conventional treatment options are exhausted? Current evidence to base recommendations upon is inadequate. The questions above need to be answered first in order to determine the degree to which slowing of disease progression and eventual disability can be attributed to treatment with new pharmacological interventions. This constitutes a high priority (and rather urgent need) for methodological research.

How has this been addressed practically. The economic evaluation by Verhoeven and associates¹¹ did not examine the longterm economic effects of the therapy. The analysis by Maetzel, *et al*¹⁵ did not model structural worsening over the 5-year time horizon and thus did not capture this potential clinical benefit in their cost-effectiveness ratio estimates. Choi, *et al*¹² modeled an exponentially increasing cost of surgery with worsening HAQ disability scores, which could be used to credit those therapies that slow worsening of HAQ scores. However, few published data exist regarding the economic and quality of life aspects of this stage of the disease. These deficiencies once again point

to the importance of a methodological research agenda in this area.

Element 12: Population risk stratification

The generalizability of economic evaluations, whether based on original trial data or on decision models, depends on how representative the model or trial population is of the target population for which decisions need to be made. Response to treatment, development of adverse events, and disease management costs are a function of disease severity and risk factor profiles among the target population that need to be adequately captured in the economic evaluation. Cost-effectiveness ratios may vary for subgroups of patients with lower or higher risk profiles. Thus economic evaluations should include a clear definition of the underlying population and clinically relevant subgroups that are at higher or lower risk of developing response or adverse consequences from the intervention. In this case a sample stratification would be more explicit than conducting sensitivity analysis.

How has this been addressed practically. None of the evaluations mentioned above specifically addressed population risk stratification as this issue may not be as pertinent in RA as in other diseases. However, response to DMARD may vary depending on whether they are given early or late in the disease course or it may depend on other biologic predictors of treatment response. However, little research exists that specifically addresses this issue and more understanding of important risk modifiers needs to be gained before incorporating these into economic evaluations.

COSTS

All published pharmacoeconomic recommendations include broad areas of resource utilization, such as costs of healthcare services, costs of patient time expended for the intervention, costs associated with caregiving, other illness-associated costs such as childcare or travel expenses, and costs associated with non-health effects of the intervention. The US Panel argued that effects of the disease on the productivity and leisure of the patient are best captured in a utility-based measure of health-related quality of life, the value of which is used in calculating QALY estimates. However, some of the indirect measures such as the HUI or the EQ-5D do not capture productivity loss explicitly.

Element 13: Which cost categories should be reported?

There is agreement that direct costs should be comprehensively assessed and accounted for in a base-case analysis and that indirect or "social productivity" costs should be reported separately. The validity of most cost assessment instruments is unknown, and further research to evaluate cost assessment instruments in rheumatology and the role and importance of indirect costs is needed.

REFERENCES

1. Menzel P, Gold MR, Nord E, Pinto-Prades JL, Richardson J, Ubel PA. Toward a broader view of values in cost-effectiveness analysis of health. *The Hast Cent Rep* 1999;29:1-10.
2. Neumann PJ, Stone PW, Chapman RH, Sandberg EA, Bell CM. The quality of reporting in published cost-utility analyses, 1976-1997. *Ann Intern Med* 2000;132:964-72.
3. Chapman RH, Stone PW, Sandberg EA, Bell C, Neumann PJ. A comprehensive league table of cost-utility ratios and a sub-table of "panel-worthy" studies. *Med Decis Making* 2000;20:451-67.
4. Canadian Coordinating Office for Health Technology Assessment. Guidelines for economic evaluation of pharmaceuticals. 2nd ed. Ottawa: CCOHTA; 1997.
5. Russell LB, Gold MR, Siegel JE, Daniels N, Weinstein MC, for the Panel on Cost-Effectiveness in Health and Medicine. The role of cost-effectiveness analysis in health and medicine. *JAMA* 1996;276:1172-7.
6. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB, for the Panel on Cost-Effectiveness in Health and Medicine. Recommendations of the Panel on Cost-Effectiveness in Health and Medicine. *JAMA* 1996;276:1253-8.
7. Siegel JE, Weinstein MC, Russell LB, Gold MR, for the Panel on Cost-Effectiveness in Health and Medicine. Recommendations for reporting cost-effectiveness analyses. *JAMA* 1996;276:1339-41.
8. National Institute for Clinical Excellence. Technical guidance for manufacturers and sponsors: new technology appraisal guidance documents. London: National Institute for Clinical Excellence; 2001:1-76.
9. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996;313:275-83.
10. Gabriel SE, Tugwell P, Drummond M. Progress towards an OMERACT-ILAR guideline for economic evaluations in rheumatology. *Ann Rheum Dis* 2002;61:370-3.
11. Verhoeven AC, Bibo JC, Boers M, Engel GL, van der Linden S. Cost-effectiveness and cost-utility of combination therapy in early rheumatoid arthritis: randomized comparison of combined step-down prednisolone, methotrexate and sulphasalazine with sulphasalazine alone. COBRA Trial Group. *Combinatietherapie Bij Reumatoïde Artritis. Br J Rheumatol* 1998;37:1102-9.
12. Choi HK, Seeger JD, Kuntz KM. A cost-effectiveness analysis of treatment options for patients with methotrexate-resistant rheumatoid arthritis. *Arthritis Rheum* 2000;43:2316-27.
13. Maetzel A, Strand V, Tugwell P, Wells G, Bombardier C. The cost effectiveness of adding leflunomide to a 5-year strategy of conventional disease modifying antirheumatic drugs in patients with RA. *Arthritis Care Res* 2002;47:655-61.
14. Boers M, Verhoeven AC, Markusse HM, et al. Randomised comparison of combined step-down prednisolone, methotrexate and sulphasalazine with sulphasalazine alone in early rheumatoid arthritis [see comments]. [erratum appears in *Lancet* 1998;351:220]. *Lancet* 1997;350:309-18.
15. Maetzel A, Wong A, Strand V, Tugwell P, Wells G, Bombardier C. Meta-analysis of treatment termination rates among rheumatoid arthritis patients receiving disease-modifying anti-rheumatic drugs. *Rheumatology Oxford* 2000;39:975-81.
16. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in health and medicine. New York, Oxford: Oxford University Press; 1996.
17. Felson DT, Anderson JJ, Boers M, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid Arthritis Clinical Trials. *Arthritis Rheum* 1993;36:729-40.
18. van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum* 1998;41:1845-50.
19. Kobelt G, Eberhardt K, Jonsson L, Jonsson B. Economic consequences of the progression of rheumatoid arthritis in Sweden. *Arthritis Rheum* 1999;42:347-56.
20. EuroQol Group. EuroQol — a new facility for the measurement of health-related quality of life. The EuroQol Group [see comments]. *Health Policy* 1990;16:199-208.
21. Furlong WJ, Feeny DH, Torrance GW, Barr RD. The Health Utilities Index (HUI) system for assessing health-related quality of life in clinical studies [review]. *Ann Med* 2001;33:375-84.
22. Woodworth TG, Furst DE, Strand V, et al. Standardizing assessment of adverse effects in rheumatology clinical trials. Status of OMERACT Toxicity Working Group March 2000: towards a common understanding of comparative toxicity/safety profiles for antirheumatic therapies. *J Rheumatol* 2001;28:1163-9.
23. Kavanaugh A, Heudebert G, Cush J, Jain R. Cost evaluation of novel therapeutics in rheumatoid arthritis (CENTRA): a decision analysis model. *Semin Arthritis Rheum* 1996;25:297-307.