Health status in rheumatoid arthritis over 7 years

A Boonen, R Landewé

The challenge is to separate fully the effect of the year of inclusion in the register and the effect of duration of the disease

A pivotal question when evaluating the long term outcome of patients with rheumatoid arthritis (RA) is how the disease itself evolves over time and to what extent new treatments contribute to such changes. Insight into the evolution of the manifestations of RA can help to define agendas for research and healthcare delivery world wide. In this issue of the *Annals* Heiberg *et al* describe “7 year changes in health status and priorities for improvement of health in patients with rheumatoid arthritis” and conclude that “although pain remained the area of highest priority for improvement among patients, the health status improved from 1994 to 2001, probably owing to access to better and more aggressive treatments”.¹ In this editorial, we will briefly elaborate on the strengths of longitudinal observational studies (LOS), in general, and on the relative importance of research on patient priorities in such studies, in particular. Next, we will discuss the subject of time and change over time in longitudinal studies in the light of the conclusion drawn from the present study.

Longitudinal observational studies and patient perspective as an outcome

The conclusion of the authors is based on an analysis of data from the Oslo RA Register (ORAR), a continuing LOS that was started in 1994 as a prevalent population based register, and annually updated with new cases since then.² Longitudinal studies are particularly helpful in answering questions on causation, prognosis, and long term outcomes in the “real world” of healthcare delivery. In comparison with randomised clinical trials (RCTs), LOS are characterised by a higher level of external validity (because patients are not selected for high disease activity or unfavourable prognosis, etc), an emphasis on long term evaluation, and a view on patients as individuals rather than as a group. A recognised disadvantage of LOS, which are uncontrolled studies, is

that you cannot distinguish the effects of therapeutic interventions from other effects, such as the natural course of disease. Although the value of LOS in answering relevant questions is undisputed, the methodology and analysis of LOS still pose several challenges. The pitfalls in interpretation of results from observational surveys were raised in 1955 in the *Annals of the Rheumatic Diseases* by Mainland,³ but it is the OMERACT initiative which should be credited with improving the standards of performing, reporting, and appraising LOS in rheumatic diseases.^{4 5}

“The OMERACT initiative has improved standards of performing, reporting, and appraising longitudinal observational studies”

Introducing the patient’s perspective as an outcome measure in LOS is definitely one of the innovations and strengths of Heiberg’s paper. The recognition that the patient’s view may diverge importantly from the classically measured health status outcomes has consequences for our approach to patients with RA, and adding a measure reflecting the patient’s perspective to the core set of domains of longitudinal and observational studies should be considered. However, the changing face of RA assessed by observational data is a more difficult problem with several traps and challenges, among which *time* in relation to disease duration, left and right *censorship* as possible confounders, and *modelling* outcome over time are the most relevant for the interpretation of the present study results. These subjects will be discussed in more detail in the next paragraphs.

Assessing change over time

Longitudinal observations inherently imply repeated measurements of variables of interest in the *same patients*. The design can apply either to a group of patients, who are sampled over a limited period (referred to as an observational

cohort), or to a dynamic group to which cases are continuously added but also removed for various reasons (usually referred to as a register). It is important to realise that in LOS (unlike in RCTs) “change over time” can refer either to increasing disease duration or to evolving calendar time. This distinction will be briefly explained.

Classically, LOS report on the evolution of the disease within patients with increasing disease duration, and often aim at identifying predictors of outcome. Ideally, the times of the repeated assessments in relation to the disease duration are similar for all patients in the study. The statistical analysis focuses on trends of the total group and within patients belonging to that group, rather than on differences between patient groups.

On the other hand, LOS also offer the possibility of examining the evolution of the disease itself over consecutive periods of calendar time. This assumes an analysis of completely separate groups (cohorts) of patients based on the calendar time of inclusion in the register, and does not necessarily require repeated measurements in the same patient. Certainly, repeated measurements within the patients of each separate cohort may add to the precision of the estimation of disease evolution, but the focus here is on differences between groups rather than on trends within one group or within patients belonging to one group. Figure 1 represents a hypothetical example of the evolution of functional disability (Health Assessment Questionnaire (HAQ)) as disease duration lengthens for two separate cohorts, one diagnosed between 1986 and 1990 and the other between 1991 and 1995.

The analysis, which might be the partially modelled data, would illustrate that in patients being diagnosed in recent years, the initial decrease in HAQ is more pronounced and the rate of worsening later on is lower.⁶ Importantly, these data have been generated by repeatedly measuring the same patients, but in *two distinct cohorts of the same register*. It shows that there is a non-linear relationship between disease duration and the modified HAQ (MHAQ), which is different according to the subgroup of period of diagnosis. This non-linearity is important; it proves that the “age of the cohort” is relevant for interpretation of the difference between MHAQ scores of both cohorts.

The question whether the impact of RA differs among separate cohorts defined by consecutive periods of diagnosis of the disease is highly relevant. However, hard evidence for this is

Figure 1 Hypothetical example of the evolution of the MHAQ in two cohorts of patients diagnosed during two different periods and further followed up longitudinally over the course of the disease.

scarce. From the ARAMIS uncompleted registry for RA the influence of the time of entry into the register on the area under the curve of the HAQ was reported for 3035 patients diagnosed between 1977 and 1998, providing 31 321 observations. A strong negative correlation ($r = -0.92$) between the year of inclusion in the cohort (calendar time) and time averaged HAQ was found. By applying generalised least squares regression (taking into account that each subject contributed more observations) and after adjusting for confounding variables such as sex, race, age, and disease duration at entry into the register the disability index declined at a rate of 2% a year (95% confidence interval (CI) 1.8 to 2.2).⁷ Mean HAQ disability at entry was 15% greater in those entering the register before 1985 than in those entering after 1994.

Left and right censorship

Heiberg suggests a similar change of the appearance of the disease.¹ Small but consistent improvements in several classical outcome measures, including physical and social functioning, pain, fatigue, and global and mental well-being are observed. This conclusion, however, is based on a quite different methodological approach. In the analysis the two concepts of time (duration of the disease and birth year of the cohort) are combined. Figure 2 shows that Heiberg's study group includes patients with two assessments (both in 1994 and in 2002) ($n = 475$), patients who only participated in the first survey but not in the second one ($n = 457$), and patients who only participated in the second survey but not in the first one ($n = 355$). By carrying out a cross sectional comparison of two populations (health status in 1994 v 2001), the longitudinal information that is locked up in their study groups is ignored. Or in other

words, by omitting the fact that part of this group was measured twice, and thus neglecting statistical dependence, the results deal with the situation as if there were two separate cohorts. This might have considerable consequences because the conclusion that—at an average disease duration of 13 years—health status in 2001 is better than in 1994 is only valid if the point estimates of 1994 and 2001 are derived from two separate cohorts or from two cross sectional populations that are intrinsically completely comparable.

Although the authors assume such comparability because their register showed high completeness at its start in 1994 (85%) and the number of patients remained fairly constant, the matter should be considered. Bias induced by left as well as right censorship needs to be considered carefully. The Oslo RA register started in 1994 as a prevalent population study, including all patients with RA known at that moment.² Some patients with RA who might potentially have been included on the basis of the average disease duration of the cohort at that moment were not sampled—for example, because they had severe disease and already died.

Figure 2 Evolution of the Oslo Rheumatoid Arthritis Register (ORAR) from its start in 1994 and the response to the 1994 and 2001 surveys.

Left censorship bias results in a prevalence cohort which did not include patients with the most severe disease, and ORAR will be no exception. Because patients included in the register after 1994 but who dropped out before the 2001 survey are not considered, the group participating in the second survey is also liable to left censorship. Importantly, there can be differences in the magnitude and causes of left censorship in both cross sectional groups. It is difficult to determine quantitatively which role left censorship played and in what direction. A subtle indication for this kind of selection process is that mean age and disease duration differ across both patient groups; the patients in the register in 1994 had a higher age and a lower disease duration than those in the 2001 register, whereas they should be similar. The authors argue this might be caused by a small decrease in the incidence of the disease,⁸⁻¹¹ explaining the lower disease duration of patients in the register in 1994, but in this case the lower mean age can only be explained if the new cases added to the register are younger at inclusion in the register. It is difficult to hypothesise whether these patients are younger at onset of disease or are diagnosed and referred earlier. Theoretically, it might also be possible that some of the new patients included after 1994 but dropping out before 2001 died because of more aggressive treatment (left censorship). Continuing registers are particularly helpful in providing information on patients who were diagnosed in between assessment periods but left the register before the next assessment, and such information should be fully exploited.

A second challenge inherent to longitudinal studies is the potential bias imposed by selective non-response (also known as right censorship bias or attrition). An indication that this kind of bias was operative in this study also, is the difference in response rate and the fact that in the first survey the non-responders were older than the responders, whereas they were younger in the second survey. The influence of attrition on results of LOS is well known, and numerous methods to adjust for attrition have been suggested.^{12 13}

One might feel that the methodological issues discussed on censorship are not relevant for the reported results, but it should be realised that the observed improvements in health status, were only minimal. Consideration of the MHAQ again as an example, shows that the disability index decreased from 1.68 in 1994 to 1.58 in 2001. The non-overlapping 95% CI confirms that this finding is statistically significant, but

should be interpreted in the light of the large number of patients included, and it should be noted that the observations were not all independent (artificially narrowing the confidence interval). Calculation of the standard deviation (SD) from the 95% CI as a measure for variability between the patients, provides an estimated SD in both groups of about 0.6 while the group difference was 0.1. The group difference of 0.014 a year on the scale of 0–3 of the HAQ is small and expressed as a relative decrease from the initial MHAQ would be 5.9% over 7 years or 0.85% a year. Certainly, small but significant improvements have been observed among all the outcomes measured in the study and add to the robustness of an overall (small) improvement.

In the light of the methodological issues considered, the pivotal question is to elucidate whether the small differences in health status over- or underestimate a possible decrease of the impact of RA over time. As explained by Heiberg, the patients who were measured twice as a group have a different change in health than the remaining patients. In the discussion the authors mention that “patients responding at both times ($n=475$) had, as expected, worse health in 2001 than in 1994 owing to the 7 year increase in age and disease duration”, implying that larger improvements must have been seen in patients who were unique to both surveys. The authors continue that “similar improvements in health status between 1994 and 2001 were seen for patients with short (<5 years) and long (>5 years) disease duration”. The exact differences in health status of these groups would have been interesting. Those with disease duration <5 years are by definition unique to each survey, whereas those with disease duration >5 years include patients with (probably) two assessment points that showed a decrease in health. Insight into these groups might help in clarifying whether the observed difference can be attributed to changing disease expression or to earlier or more aggressive treatment, as suggested in the article.

Modelling outcome over time

The question whether RA changes over time (whatever the reason: disease expression, earlier diagnosis, better treatment) is challenging and relevant, but providing the right answer is difficult. Although the theoretical model suggested by the authors of this paper in fig 1 has high face validity, it is merely used as an example to show clearly the difference between the “time” referring to disease duration and “time” referring to the inclusion in the cohort. We realise that such an analysis is not feasible. The large number of patients needed for each cohort and (very) long observations for each cohort make it unlikely that most continuing registers can ever be analysed by such an approach. In the study by Krishnan and Fries calendar time at entry into the register (while correcting for age and disease duration at entry) was used as an independent variable in multivariate regression as a surrogate for assessing the expression of the disease (area under the curve) over time.⁷ As always, all methods have advantages and disadvantages and the best choice is influenced by the question and data at hand.

CONCLUSION

In conclusion, the paper by Turid Heiberg *et al* states that the patients’ perspective is an important outcome in LOS but, in the true spirit of the scientific quest to obtain correct answers to hard questions, allows for the possibility of some criticism. The term cohort should be reserved for distinct groups of patients based on the period of inclusion in the study. Continuing registers can potentially reduce left censorship, but this implies that all patients who were (even if only for a short period) part of the register will be accounted for in the longitudinal analysis. When reporting results from continuing registers it is a challenge to separate fully the effect of the year of inclusion in the register and the effect of the duration of the disease.

Ann Rheum Dis 2005;**64**:173–175.
doi: 10.1136/ard.2004.033407

Authors’ affiliations

A Boonen, R Landewé, Department of Internal Medicine, Division of Rheumatology, University Hospital Maastricht and Caphri Research Institute, The Netherlands

Correspondence to: Dr A Boonen, Department of Internal Medicine, Division of Rheumatology, University Hospital Maastricht, PO Box 5800, 6202 AZ Maastricht, The Netherlands; aboo@sint.azm.nl

REFERENCES

- 1 Heiberg T, Finset A, Uhlig T, Kvien TK. Seven year changes in health status and priorities for improvement of health in patients with rheumatoid arthritis. *Ann Rheum Dis* 2005;**64**:191–5.
- 2 Kvien TK, Glennas A, Knudsrød OG, Smedstad LM, Mowinckel P, Forre O. The prevalence and severity of rheumatoid arthritis in Oslo. Results from a county register and a population survey. *Scand J Rheumatol* 1997;**26**:412–18.
- 3 Mainland D. Use of case records in the study of therapy and other features in chronic disease. *Ann Rheum Dis* 1955;**14**:337–51.
- 4 Wolfe F. Critical issues in longitudinal and observational studies: purpose, short versus long term, selection of study instruments, methods, outcomes, and biases. *J Rheumatol* 1999;**26**:469–72.
- 5 Wolfe F, Lassere M, van der Heijde D, Stucki G, Suarez-Almazor M, Pincus T, *et al*. Preliminary core set of domains and reporting requirements for longitudinal observational studies in rheumatology. *J Rheumatol* 1999;**26**:484–9.
- 6 Wolfe F. A reappraisal of HAQ disability in rheumatoid arthritis. *Arthritis Rheum* 2000;**43**:2751–61.
- 7 Krishnan E, Fries JF. Reduction in long-term functional disability in rheumatoid arthritis from 1977 to 1998: a longitudinal study of 3035 patients. *Am J Med* 2003;**11**:371–6.
- 8 Jacobsson LT, Hanson RL, Knowler WC, Pillemer S, Pettitt DJ, McCance DR, *et al*. Decreasing incidence and prevalence of rheumatoid arthritis in Pima Indians over a twenty-five-year period. *Arthritis Rheum* 1994;**37**:1158–65.
- 9 Shichikawa K, Inoue K, Hirota S, Maeda A, Ota H, Kimura M, *et al*. Changes in the incidence and prevalence of rheumatoid arthritis in Kamitonda, Wakayama, Japan, 1965–1996. *Ann Rheum Dis* 1999;**58**:751–6.
- 10 Kaipiainen-Seppänen O, Aho K, Isomaki H, Laakso M. Incidence of rheumatoid arthritis in Finland during 1980–1990. *Ann Rheum Dis* 1996;**55**:608–11.
- 11 Linos A, Worthington JW, O’Fallon WM, Kurland LT. The epidemiology of rheumatoid arthritis in Rochester, Minnesota: a study of incidence, prevalence, and mortality. *Am J Epidemiol* 1980;**111**:87–98.
- 12 Krishnan E, Murtagh K, Bruce B, Cline D, Singh G, Fries JF. Attrition bias in rheumatoid arthritis databanks: a case study of 6346 patients in 11 databanks and 65,649 administrations of the Health Assessment Questionnaire. *J Rheumatol* 2004;**31**:1320–6.
- 13 Farewell VT. Studies of attrition in rheumatological databases. *J Rheumatol* 2004;**31**:1244–5.