Reporting Requirements for Longitudinal Observational Studies in Rheumatology

The following is a list of recommendations of core methodological items that should be included in all publications on longitudinal observational studies in rheumatology. Information about these is essential for both appraisal of quality and incorporation into metaanalyses.

CORE METHODOLOGICAL ITEMS

Longitudinal observational studies should include the following core items.

1. Study design type: true prospective, retrospective, or mixed.
2. Source of cases: true population-based, catchment population, consecutive series (specify clinic type), or other.
3. Timing of patient recruitment in relation to disease onset (to enable estimation of left censorship bias): cases followed from disease onset, cases followed from first presentation, or prevalent cases.
4. Inclusion criteria: classification criteria, age range, sex.
5. Demographic data collected: sex, age, socioeconomic factors, ethnic group.
6. Baseline clinical data collected. Specify individual items of data collected at baseline. Distinguish between items ascertained from routine medical records (errors or missing data probable) and items collected prospectively using a standard proforma. Specify number of observers, training requirements, and any measures of observer variability.
7. Followup data collection. Specify frequency of followup and decision rules about timing of each assessment. Provide information on proportion of patients with missing followup information at each individual time point and estimate potential for loss to followup bias (right censorship). Indicate means of followup data collection (clinical interview, questionnaire, mail or telephone). Report number of observers involved in prospective data collection, nature of training, and report on observer variability. Report on the principal and subsidiary outcome measures chosen. Comment on observer blindness to baseline variables.
8. Analysis: specify strategies used for missing data and loss to followup. Indicate, in relation to person-years of followup, the power to detect clinically meaningful differences for the major outcomes analyzed. If a statistical model is generated, indicate performance in a validation sample.

RATIONALE

Well conducted longitudinal studies provide valuable information on the cause and outcome of such disorders and may also contribute to knowledge about the relative effectiveness of different interventions in situations where randomized, prospective clinical trials are impractical. It is necessary to standardize reporting requirements for 2 reasons: (1) overviews and formal statistical metaanalyses require studies to be comparable so that data can be pooled; (2) the contribution of a study to the body of knowledge on outcome can only be considered in relation to the perceived quality of the investigation. The availability of data on these main methodological items will inform such quality judgment. As an additional benefit, potential investigators might improve the quality of their investigations by attention to these details at the protocol development stage.

STUDY DESIGN: SPECIFY WHETHER TRUE PROSPECTIVE, RETROSPECTIVE, OR MIXED

In true prospective studies, the investigator initiates the baseline data collection at the start of the followup period. This has the advantage of ensuring quality in both baseline and followup data collection and the development of strategies to ensure maximal participation at followup. Many longitudinal studies are in fact retrospective, based on medical record (chart) review of subjects attending a clinical facility in the past. While this has the advantage of speed, there is also a cost, since the quality of data collected during routine clinical practice may not be sufficient in all cases or for all data items to provide valid information. Some studies may be mixed: for example, the investigator initiates the data collection at onset in some subjects and incorporates retrospective data from others.

COHORT SELECTION: SPECIFY WHETHER TRUE POPULATION BASED, CATCHMENT POPULATION, CONSECUTIVE SERIES FROM A SPECIFIED CLINIC, OR OTHER NON-RANDOM SERIES OF PATIENTS

The ideal, but infrequently undertaken, study is to include all patients with the disorder under investigation arising from a specified population. This requires that there is a case ascertainment scheme in progress that will capture all cases. For most studies, recruitment is restricted to patients attending a rheumatologist — this will exclude subjects seen by a generalist and those who may attempt self-care or seek care from heterodox practitioners. The second best option is the "catchment population approach," where the aim is to identify from clinical facilities all cases that have arisen from a fixed catchment population, normally geographical, but possibly administrative. Such an approach necessitates that subjects referred from outside the population to the same

clinical facilities be excluded. Conversely, patients from within the catchment population that have sought health care outside the normal clinic facilities should be captured. Inclusion of the former would lead to error, as they are likely to include selectively those with severe disease referred for a specialist opinion. As an example, longitudinal survival studies of the relatively rare connective tissue diseases, such as scleroderma and SLE, frequently come from tertiary referral centers. Patients attending these centers are a combination of local patients and those referred from outside the area for whom, possibly, there is a particular diagnostic or severity problem. Clearly, their survival experience would not be representative of the survival experience of all patients with the disorder.

Often, a consecutive series of clinic attenders is included as the population base for study. The nature of the clinic should be specified and the likely sources of bias that this introduces. As an example, patients with low back pain attend a large number of different clinical services within a health service institution. It may be envisaged, therefore, that subjects with back pain attending orthopedic surgeons, neurologists, neurosurgeons, or rheumatologists may have substantial differences, both in baseline characteristics and outcome. Investigators must be honest as to whether the patient series was genuinely a true consecutive series or, as is frequently the case, a haphazard series of patients on whom there is available data or who have agreed to attend regularly for followup.

PROXIMITY OF RECRUITMENT TO TRUE ONSET OF DISEASE (AVOIDANCE OF LEFT CENSORSHIP BIAS)

Information on disease duration should be collected in all longitudinal studies. Left censorship in longitudinal studies refers to the potential bias introduced when patients are recruited at some stage after disease onset. The concern is that other patients who developed the disease at the same time may be excluded, either because they had very severe disease and did not survive long enough to be included, or because they had mild disease that had resolved by that equivalent time point. The only way to avoid this bias is to ensure that information is available on all the eligible patients from the time of first onset of disease. It has to be accepted that with many of the rheumatic disorders this is very difficult. For example, in the seronegative spondyloarthritides, there may be a considerable delay between first recorded symptom such as low back pain and first presentation to a medical practitioner. The alternative therefore is to base recruitment on subjects' first attendance. This does not exclude left censorship, but it allows the reader to make judgments about the nature of the referred population. The least desirable choice is to study a prevalent cohort, i.e., select only those patients who are current attenders — this exaggerates the likelihood of left censorship bias, and inferences about outcome from such a group only apply to other groups constituted in exactly the same way.

SPECIFY INCLUSION CRITERIA

Comparative studies require the use of standardized classification criteria, which, for most rheumatic disorders, are the relevant criteria published under the aegis of the American College of Rheumatology. It is important, particularly in retrospective studies relying on review of medical charts, to indicate the proportion of subjects for whom it is impossible to verify with 100% accuracy the classification status. It is preferable to include a group where the classification status is indeterminate due to missing data, rather than to rely solely on those subjects for whom there are sufficient data for that purpose. The latter approach may lead to selection bias. As an example, in rheumatoid arthritis, medical record information on the presence or absence of erosions is far more robust than data on symmetrical joint involvement. Over-reliance on the former, however, would lead to a more severe cohort being included. Other inclusion criteria that should be specified include age extremes and whether other factors were involved, for example, language or other skills necessary to participate in the followup process, whether the recruitment was restricted to specific ethnic or socioeconomic groups or to individuals under particular health payment plans.

SPECIFY DETAILS OF BASELINE DATA COLLECTION

The essence of the analytical approach to longitudinal studies is the relationship of specific items of baseline data to outcomes. The quality of the information collected at baseline is therefore of crucial importance. It is important therefore to specify those items that are collected prospectively and (see above) distinguished from those items gathered from medical chart reviews. For items of clinical data requiring observer interpretation, it is important to indicate the number of observers, whether observer training for reliability was undertaken, and whether, over prolonged periods of recruitment, reliability was checked over the same time period. The number of individuals with missing information for each of the collected items should also be stated. Wherever possible, validated and published measures should be used. It is not appropriate to develop and then use a new instrument in the same study.

SPECIFY DETAILS OF FOLLOWUP DATA COLLECTION

Most prospective studies will have varying lengths of followup for the individual subjects recruited. This inevitably leads to the problem of right censorship, i.e., at a given point in time, only a proportion of individuals who will ultimately be affected will have developed the outcome of interest. In addition, there will be missing data up to that time point

because of mortality, loss to followup, withdrawal, migration, etc. It is crucial, therefore, to ascertain the reasons for loss to followup and to compare the baseline and other interval characteristics between those remaining and those excluded from the cohort. Followup may be made by individual visit, interview over the telephone, mailed questionnaire, or other means. Where different approaches are used for different subjects, this should be specified. For example, it is not an unreasonable strategy to have individuals attending clinic assessed for disability using a questionnaire administered by a research nurse. It may be decided to send individuals lost to followup a questionnaire by post. This should be clearly stated and the possibilities that the different methods of data collection may lead to different results should be explored. As with the baseline data, information about the number of individuals and the number of items for which there is missing data at followup should be specified, with some comparison of baseline characteristics for those with and without that missing information in order to assess the potential for loss-to-followup bias. Again, the number of observers involved in followup and their training and reliability should be stated. Further, the blindness of the observers to the baseline status of patients, and to the major hypotheses under test, should also be stated. As in the baseline situation, validated measures should be used wherever possible.

SPECIFY DETAILS OF THE ANALYSIS

Details of statistical analysis should be provided in the same way as for other study designs such as randomized clinical trials. In longitudinal observational studies, one key element of the analysis is the influence of baseline or interval predictor variables on subsequent outcome. These might be influenced by confounders including age and sex. Given the sample size, which should normally be expressed in person-years, the power to detect specific strengths of association of baseline predictors for particular outcomes should be stated. It is often appropriate to use multivariate modelling techniques to determine accurately the relationship between baseline predictors and outcome. The assumptions behind these models should be stated and tested. As an example, age may often be entered as a continuous variable, whereas the influence of age may be nonlinear. It should be recognized that models generated from one dataset about the value of the baseline predictors in explaining outcome may not be extrapolatable to other datasets. The purpose of statistical modelling is to provide the best model for that one set. An appropriate statistical method should, therefore, be used and specified for validating this model in a second dataset. A typical approach is to generate the model from a random three-fifths of the dataset (the test sample) and to test the robustness of that model in the remaining two-fifths. With small numbers of subjects, other statistical techniques (e.g., "jack-knifing") may be used.

ALAN SILMAN, MD, FRCP,
Director;
E-mail:Alan.Silman@man.ac.uk;
DEBORAH SYMMONS, MD, FRCP,
Deputy Director,
ARC Epidemiology Research Unit,
University of Manchester,
Manchester, UK M13 9PT.

Address reprint requests to Professor Silman.