BACKGROUND

In 1990 a body of surgeons decided to set up a user group to share their experiences of the use of a common type of knee arthroplasty (TKR) (pp 185-90 of this issue). We wished to collect sound information which would help surgeons to assess their use of a particular implant, to monitor it for early problems and to offer long-term surveillance. We agreed to meet annually to share information and indicate difficulties, thus auditing our own work. The project was funded by industry, but the data were handled independently in a single centre, which allowed identification of patients by the surgeons only. A number of problems were encountered in organising the project and setting up the databases. We consider, however, that there are mutual benefits to the patients, the surgeons and to industry in working in partnership to produce such registries.

WHAT INFORMATION SHOULD BE COLLECTED?

We used the American Knee Society Score (AKSS) to assess our patients and this has proved satisfactory, although there have been a number of problems. At the beginning of the study, many surgeons expressed preferences for collecting specific information which addressed their own interests and many wished to set a broad agenda. A compromise was eventually struck in that some additional data were collected, in particular to see whether we could detect a true incidence of deep-vein thrombosis (DVT). We were also persuaded to leave space for free-text responses so that special comments could be made which, in the view of many surgeons, extended the scope of the study. This decision proved to be wrong. A surgeon with no complications had no need to use the free text, but frequently did, which made it difficult for lay data processors to interpret the information. Medical complications were intermixed and difficult to construe. In general, there was low reporting of adverse events. We now simply ask “Were there any technical adverse effects, Yes or No?” In future we will check the operation note of the handful of affirmative answers, rather than waste time analysing free-text responses. Attempts to collect information concerning DVT and pulmonary embolus produced material which was clearly inaccurate and of anecdotal interest only. We no longer collect such data.

A considerable amount of information was accumulated both at the time of admission and at operation, but much of it was unreliable and a waste of resources. The more details required the less compliant were the surgeons. Our experience indicates that the information requested should be kept to a minimum.

We were also convinced of the value of paper collection with separate electronic data entry. Direct electronic entry tends to reinforce the vagueness of free-text entry and limits the ability of the data manager to control this area of the work. The database becomes clogged with material of poor quality which is essentially meaningless. We are now looking again at direct electronic entry in order to collect a smaller and better-defined data set.

THE EFFECTS OF MISSING DATA

The results expressed are generally in terms of means. It was noticeable, however, that unlike the preoperative data, the two- and five-year results have differences in our study between the mean and median, which suggest that the distribution is skewed. The medians are generally higher than the means but, more interestingly, the modal value is commonly higher, suggesting that failures do bring down the average score. This reinforces the point made by Murray, Carr and Bulstrode that failures must be taken into account and not be selectively ignored when considering long-term survivorship. However, follow-up revealed that, of the defaulters, some had died and that most had not replied either due to comorbidities or because they were well and had no problem with their knee. None declined follow-up because they had problems with their knee.

THE AMERICAN KNEE SOCIETY SCORE

Of more concern, was the analysis of the different domains within the AKSS, which reinforced concerns about the tendency to treat scoring systems as a set of continuous variables.
Reliability analysis was carried out on the components of the AKSS. Cronbach’s alpha\(^2\) (a model of internal consistency based on the average interitem correlation) was calculated to assess the internal consistency of the scoring system. The AKSS was found to have a Cronbach’s alpha of 0.003, which is extremely low, and indicates that the component parts of the scale are not related. This was expected since there is a subtractive element in the AKSS and the only reliable measures are based on a purely additive scale. The function section yielded a Cronbach’s alpha of 0.2389 and is low for the same reasons as for the knee score.

In an attempt to assess the qualities of the AKSS further, the purely additive components of each section were analysed. The positive section comprising pain, range of movement, anteroposterior stability and medial/lateral stability gave a Cronbach’s alpha of 0.1552 which is higher than the whole scale but still very low. The negative section comprising flexion contracture, alignment and extensor lag gave a Cronbach’s alpha of 0.03.

The positive section of the function score comprising walking distance and ability on stairs gave a Cronbach’s alpha of 0.69, which was not surprisingly the best result. The negative section of the function score could not be analysed since it consisted of only one component, namely walking aids.

The definition of a good outcome for a TKR is elusive. Most surgeons and patients recognise what is good or excellent, but expressing this objectively and repeatedly when observed by the same and different observers over time is exceptionally difficult. Attempts to do this have been well expressed in a recent series of Instructional Course Lectures from the American Academy of Orthopaedic Surgeons.\(^3\) The resulting project, sponsored by the AAOS, has been very costly and the outcome criteria are comprehensive, but it is difficult to envisage that the resources needed to complete this exercise consistently and effectively are likely to be universally available. More pragmatic approaches include the Oxford Knee Score\(^4\) and generic arthritis scores such as the Western Ontario and McMaster arthritis scoring system.\(^5,6\) It is now appreciated that any scoring system, whether implant- or disease-specific, must be supported by entirely generic information on health status. The most universally acclaimed is the internationally validated Short Form 36\(^7\) and its pragmatic offshoot the SF12.\(^8\) These methods of assessment of outcome emphasise the necessity for retaining the patient at the centre of the measure.

A COST ANALYSIS OF THE TRIAL

Costs were largely related to staff salaries with a modest amount for the initial setting-up. In some centres staff were employed and in others consultants undertook the task, or delegated it to junior and secretarial staff while retaining control; their time should be taken into account in any detailed analysis. However, allowing for the first-year costs including setting-up, and the ten-year accumulated costs in the co-ordinating centres with a ten-year inflation factor, the overall expense was modest, especially since it included the central co-ordination of 15 centres, giving support, quality assurance, feedback and digital storage of radiographs. The ultimate accumulated cost per patient of collecting data over ten years is £261 or £26 per annum. This seems to be a very reasonable price to pay for gathering accurate information.

CONCLUSIONS

The experience gained in the organisation of the IBII user group has proved to be invaluable in setting up new groups, designing trials and carrying out audit. The limitation of existing scoring systems have been highlighted, especially as a predictive tool. The work has demonstrated that it is possible to collect data reliably with a high degree of compliance. We found no undue influence from industry. There have been, and must not be, any preconditions set in such contractual arrangements concerning the rights of editing of material by industry. The independence of the data centre in scientific terms is important and the use of independent data recorders who have a vested interest in accuracy and completeness of the data rather than in the outcome of the study, cannot be overemphasised.

The IBII user group were funded by Zimmer UK in the form of an ongoing departmental research grant.

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

References

5. WOMAC. Osteoarthritis Index – a user’s guide, Nicholas Bellamy.