Predictors of hip joint replacement in new attenders in primary care with hip pain

Fraser Birrell, Cara Afzal, Elizabeth Nahit, Mark Lunt, Gary J Macfarlane, Cyrus Cooper, Peter R Croft, Gillian Hosie and Alan J Silman

SUMMARY

Background: *Studies investigating the factors associated with need for total hip replacement should ideally be based on prospective investigation of new attenders in primary care.*

Aim: *To determine the incidence of listing for total hip replacement, and its predictors, among attenders in primary care with a new episode of hip pain.*

Design of study: *Prospective multicentre cohort study.*

Setting: *One hundred and ninety-five patients (mean age = 63 years, 68% female) with new episode of hip pain, attending primary care between November 1994 and October 1997. At the first visit, patients were evaluated for indices of pain and disability, range of hip movement, and radiographic changes of osteoarthritis.*

Method: *General practitioner participants were recruited from the membership of the Primary Care Rheumatology Society to recruit all consecutive attenders with a new episode of hip pain. Annual follow-up was carried out to determine which patients were being 'put on a waiting list' for total hip replacement.*

Results: *Seven per cent of patients were put on a waiting list for total hip replacement within 12 months and 23% of patients within four years. At presentation, pain duration, pain severity, (including the need to use a stick) and restriction of internal rotation were the major clinical predictors of being put on a waiting list. Radiographic predictors of osteoarthritis performed similarly to the clinical measures. A simple scoring system based on both radiographic severity and two of the clinical measures was derived that identified groups at high likelihood of being put on a waiting list (sensitivity = 76%) with a low false-positive rate (specificity = 95%).*

Conclusion: *New primary care attenders with pain are frequently accepted for total hip replacement soon after their first attendance — a decision that can be predicted by simple clinical measures.*

Keywords: *total hip replacement; waiting list; orthopaedic referral; primary care.*

F Birrell, MD, FRCP, clinical research fellow; C Afzal, BSc, research assistant; E Nahit, PhD, research assistant; M Lunt, MSc, lecturer in biostatistics; A J Silman, MD, FRCP, professor of rheumatic disease epidemiology, ARC Epidemiology Unit; G J Macfarlane, PhD, MB ChB, professor of epidemiology, ARC Epidemiology Unit and Unit of Chronic Disease Epidemiology, University of Manchester. C Cooper, PhD, FRCP, FMedSci, professor of rheumatology, MRC Environmental Epidemiology Unit, University of Southampton; P R Croft, MD, MRCGP, professor of primary care epidemiology, Centre for Primary Care Sciences, University of Keele; G Hosie, FRCP, general practitioner, Great Western Health Centre, Glasgow.

Address for correspondence

Professor Alan J Silman, ARC Epidemiology Unit, University of Manchester Medical School, Manchester M13 9PT. E-mail: alan.silman@man.ac.uk

Submitted: 26 March 2002; Editor's response: 17 June 2002; final acceptance: 28 August 2002.

Introduction

TOTAL hip replacement is one of the most frequently undertaken surgical procedures in orthopaedic practice, with osteoarthritis of the hip being the major indicator. In the United Kingdom in 1996, there were an estimated 46 000 operations carried out and this number is set to increase substantially owing to demographic change.[1] Pain and disability are the most important indicators for surgical intervention. In patients accepted for arthroplasty there is considerable evidence of structural damage.[2] In populations, however, there is only relatively weak concordance between the amount of pain and the degree of radiographic change.[3] The demand for total hip replacement is determined by the frequency of orthopaedic referral, but this does not necessarily reflect the underlying need. Health service planning thus requires understanding of the natural history following new attendance with hip pain, but this is difficult to undertake in secondary care given the wide variability in referral and attendance patterns. Thus, studies based in primary care might provide greater insight into the potential burden. We have recently shown[4] that, depending on the radiographic criteria used, 30% to 35% of all patients attending with a new episode of hip pain in primary care over the age of 40 years already have substantive evidence of radiographic damage. In this report we have assessed the likelihood that attendance with a new episode of hip pain in primary care results in acceptance for total hip replacement in the short-term (within four years of first attendance). We also aimed to identify what are the predictors around the time of presentation to primary care with a new episode of hip pain and of subsequently being put on a waiting list for hip replacement.

Method

Design

The study was a prospective cohort study examining the likelihood of being put on a waiting list for total hip replacement among attenders in primary care with a new episode of hip pain.

Patients

Details of recruitment of these patients have been described in detail elsewhere.[4-6] In brief, general practitioner (GP) participants were recruited from the membership of the Primary Care Rheumatology Society (see Acknowledgements): a group of GPs with a specific interest in rheumatology. The participants agreed to recruit all consecutive attenders with a new episode of hip pain. This was defined as attendance with hip pain which, on a pain diagram, was confined to the bathing trunk area and which, in the GP's clinical opinion,

HOW THIS FITS IN

What do we know?

Patients with a new episode of hip pain frequently have radiographic damage at presentation. The degree of pain and disability is the major factor influencing acceptance for joint replacement.

What does this paper add?

Approximately one-quarter of all new hip pain attenders are accepted for joint replacement within four years. Easily measured clinical variables, as well as radiographic assessment at presentation, can accurately predict those thus accepted.

was arising from structures in the hip joint. Patients were excluded if they had had any previous attendances with that same episode of hip pain or if they had attended with any episode of hip pain within the previous 12-month period.

Baseline assessment

All patients completed a self-administered questionnaire providing details on the duration and severity, using a visual analogue scale, of their pain and its impact on function and quality of life, including the completion of the Short Form 36 (SF36).[7] The SF36 findings are expressed as Z-scores for each domain, as described elsewhere,[5] to allow for the influence of age and sex on these findings. A Z-score of −1.0 indicates that, for that domain, the subject's value is one standard deviation below that expected for the population of that age and sex. The patients were examined by the GPs and the range of movements were examined in three planes: internal and external rotation, and flexion, following a standardised procedure described elsewhere.[8] All patients were requested to attend their local radiology department for a standardised anterior posterior pelvic radiograph. These films were read by two trained observers using (a) the Croft modification[9] of the Kellgren and Lawrence grading scheme for hip osteoarthritis, and (b) minimum joint space. Both inter- and intra-observer reliability with these observers was found to be good.[4] Discrepancies of one grade or minimum joint space of greater than 0.5 mm were adjudicated by a third independent observer.

Follow-up

Patients were followed-up annually both by postal questionnaire and by contact with the patients' GPs to determine (a) whether or not they had been referred for orthopaedic opinion, and (b) whether or not they had been put on a waiting list for total hip replacement.

Statistical analysis

A survival analysis approach was adopted using months of observation from the time of recruitment until the study endpoint. The key outcome was the date of being put on a waiting list for total hip replacement. Follow-up was continued to the end of December 1999, or earlier for those patients who had died, or who had moved elsewhere and could not be contacted. Cox proportional hazards modelling was undertaken to determine the role of the baseline clinical, examination, and radiographic variables on the likelihood of being put on a waiting list. Initial univariate analysis was conducted, adjusting for age and sex. Variables that emerged that were significantly associated with being put on a waiting list were then entered as candidate variables into a forward stepwise regression model. A Wald test was undertaken to assess the statistical significance of the improvement in the model from adding the different categorical variables evaluated. Two models were evaluated: (a) clinical predictors excluding radiographic findings; and (b) clinical predictors and radiographic findings. The fit of each model in predicting the likelihood of being put on a waiting list was assessed using Harrell's 'c' statistic.[10] This is equivalent to the area under a Receiver Operating Characteristic (ROC) curve but allows for variable follow-up time. The statistic can be interpreted as the probability that, of two randomly chosen patients, the patient with the higher score will be put on a waiting list for surgery before the patient with the lower score.

Results

In all, 195 patients were recruited between November 1994 and October 1997. Their mean age was 63 years (SD = 11 years) and 132 (68%) were female. Most (190 [97%]) had unilateral pain and only 74 (38%) reported any previous episode of hip pain.

The key clinical and radiographic findings at first attendance of these patients are shown in Table 1. Not surprisingly, these patients recorded high median pain severity based on a visual analogue scale. The patients displayed a reduction in the key aspects of physical functioning and physical role in relation to the SF36, as well as pain as discussed above (Figure 1) and as previously reported.[4] Many of these patients already had substantial evidence of radiographic damage on first presentation, with 59 (30%) having

Table 1. Descriptive data for cohort.

Feature	
Number of patients	195
Female sex *n* (%)	132 (68)
Mean age in years (SD)	63 (11)
Side of pain	
Right *n* (%)	103 (53)
Left *n* (%)	87 (45)
Both *n* (%)	5 (3)
Range of movement	
Flexion: median (IQR)	98° (84°–110°)
Internal rotation: median (IQR)	26° (20°–34°)
External rotation: median (IQR)	28° (20°–36°)
Pain VAS score: median (IQR)	5 (3–7)
Pain duration[a]	
Less than 3 months *n* (%)	58 (30)
3–12 months *n* (%)	80 (41)
More than 12 months *n* (%)	54 (28)
Reported previous episodes of hip pain *n* (%)	74 (38)

[a]Pain duration was not available for three patients, hence the total is not equal to 100%. VAS = visual analogue scale; IQR = interquartile range.

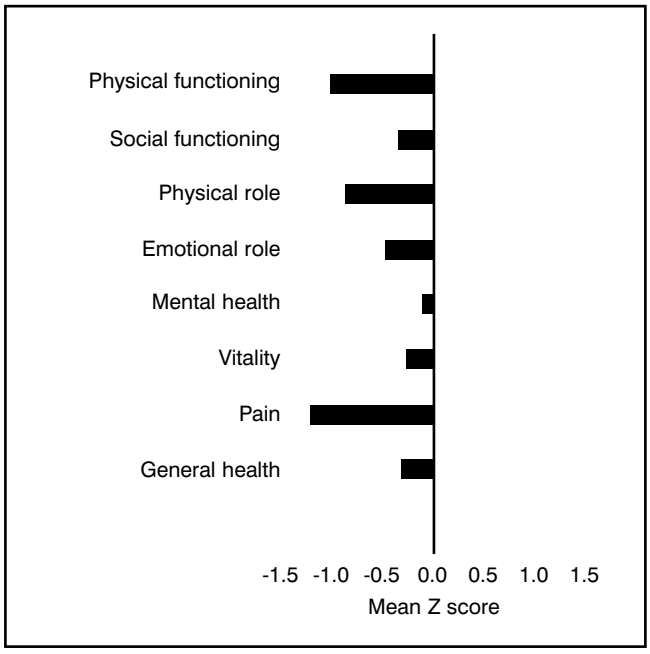

Figure 1. Baseline SF36 scores, adjusted for age and sex.

a minimum joint space of less than 2.5 mm and 67 (34%) having a Croft Grade of greater or lesser than 3.

The median duration of follow-up in the cohort was 36 months (IQR = 24 to 48 months). In all, 22 (11%) were lost to follow-up, comprising nine deaths, seven moves, four severe illnesses, and only two refusals. During follow-up, 33 (16%) were put on a waiting list for hip joint replacement, and all these patients have actually received that surgery. The Kaplan–Meier survival plot of the time to being put on a waiting list is shown in Figure 2. In all, 7% of patients had been put on a waiting list by the first anniversary of attendance and this figure had increased to 23% by four years.

The univariate predictors of being put on a waiting list for hip replacement are shown in Table 2, expressed as hazard ratios with 95% confidence intervals. As expected, pain score on the visual analogue scale was a powerful predictor, with a 25% increase in risk per unit change in score. Interestingly, severe restriction in movement, particularly of internal rotation, was the best of the clinical predictors. In addition, other indicators of pain severity, including the use of painkillers prior to first attendance, and markers of disability, including the use of a walking stick and interference with function, particularly physical function, were all associated with an increased likelihood of being put on a waiting list. Using a forward stepwise selection, restriction in internal rotation, pain score, pain duration, and use of a walking stick were all independent predictors. The 'c' statistic for a model based only on those predictors was 0.82. This increased only marginally to 0.84 when all clinical predictors were allowed to enter the model.

Separate models were also derived, based only on the radiographic data. These items were almost identical in their prediction, with 'c' statistics for minimum joint space of 0.82 and for a Croft score of 0.85.

All variables, both clinical and radiographic, were then entered into a Cox regression model, which generated an

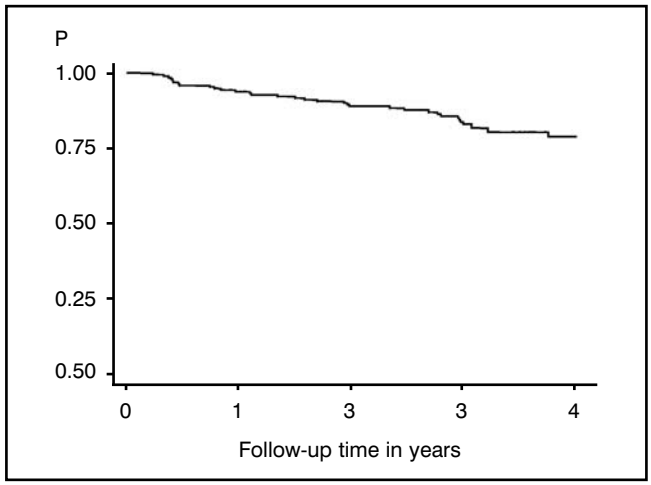

Figure 2. Kaplan–Meier plot of time to being put on a waiting list.

overall score of six points (Table 3) derived from just four variables: both radiographic items, the use of a walking stick, and the pain score, all of which were independent predictors. The results of applying this model (Table 3) show that those with a low score (less than two points) had only a 2% chance of being subsequently put on a waiting list, compared with one of 88% for those with a high score (more than five points). A score of three points or greater had a sensitivity of 76% with a specificity of 95%, equivalent to a likelihood ratio of 15, a potential useful cut-off for clinical practice.

Discussion

Summary of main findings

In summary, approximately a quarter of new attenders to primary care with hip pain, in this cohort, were put on a waiting list for total hip replacement within four years of being seen. Severe pain, disability, and restriction in movement, particularly internal rotation, proved important predictors. In addition, radiographic evidence of osteoarthritis proved of equal value in identifying those patients who had subsequently been put on a waiting list. A combined simply derived score, based on four of these predictors, discriminated between those patients who were subsequently put on a waiting list and those who were not.

This study has demonstrated the relative contribution of clinical and radiographic variables in statistically predicting the decision that a patient will be referred for arthroplasty. Indeed, clinical and radiographic predictors on their own are similar in their performance, although in combination (as shown in Table 3) they provide additional discrimination.

Strengths and limitations

There are a number of limitations of this study that need to be considered when interpreting the results. First, the external validity in relation to the patients recruited should be addressed. As with patients recruited from secondary or tertiary care centres, patients attending primary care with hip pain may vary in their disease severity, both within and between GPs. We attempted to overcome some of these concerns by recruiting from multiple GPs. The members of

Table 2. Univariate predictors of being put on a waiting list.

Predictor	Category	Hazard ratio (95% CI)	*P*-value (Wald test)
Clinical			
Pain score	Per unit	1.25 (1.07–1.46)	0.006
Pain duration (months)	<3	1 (referent)	0.09
	3–12	1.30 (0.51–3.31)	
	>12	2.49 (1.00-6.16)	
Use of analgesics	No	1 (referent)	0.06
	Yes	2.77 (0.97–7.87)	
Use of walking stick	No	1 (referent)	<0.0001
	Yes	5.22 (2.48–11.03)	
SF36 physical functioning Z-score	Per unit	0.80 (0.63-1.02)	0.08
Range of flexion (quintile)	Highest	1 (referent)	0.03
	2nd	1.39 (0.23–8.35)	
	3rd	2.06 (0.40–10.61)	
	4th	3.19 (0.66–15.39)	
	Lowest	7.59 (1.72–33.43)	
Range of internal rotation (quintile)	Highest	1 (referent)	<0.0001
	2nd	0.50 (0.04–9.43)	
	3rd	7.79 (0.96–63.37)	
	4th	7.15 (2.89–57.13)	
	Lowest	16.69 (2.20–126.41)	
Range of external rotation (quintile)	Highest	1 (referent)	0.02
	2nd	0.83 (0.17–4.13)	
	3rd	1.18 (0.26–5.27)	
	4th	3.22 (0.91–1.42)	
	Lowest	3.62 (0.99–13.14)	
Radiographic			
Croft grade	0/1	1 (referent)	<0.0001
	2	3.36 (0.31–38.19)	
	3	15.23 (3.29–70.49)	
	4	44.51 (10.04–197.48)	
	5	57.29 (12.12–270.71)	

Table 3. Prediction for being put on a waiting list for hip replacement: influence of combined radiographic and clinical score.

Score[a]	Number	Put on a waiting list *n* (%)	Sensitivity (%)	Specificity (%)
0	53	1 (1.9)	100	32.1
1	57	1 (1.8)	97	66.7
2	38	6 (15.8)	93.9	86.4
3	17	3 (17.6)	75.8	95.1
4	13	7 (53.8)	66.7	98.8
5	9	7 (77.8)	45.5	100.0
6	8	8 (100.0)	24.2	100.0

[a]Score calculated on following four variables: minimum joint space >2.5 mm = 0; minimum joint space 1.5–2.5 mm = 1; minimum joint space <1.5 mm = 2. Croft grade 0–1 = 0; Croft grade 2–3 = 1; Croft grade 4–5 = 2. Pain Visual Analogue Scale (VAS) <5 mm = 0; pain VAS 5 mm = 1. Use of walking stick 'no' = 0; use of walking stick 'yes' = 1.

the Primary Care Rheumatology Society were, however, volunteers and enthusiastic about the study, and their referral policy for patients with hip pain may not represent the referral patterns of the broad mix of GPs elsewhere in the UK. The GP participants were also asked to recruit all consecutive new attenders with hip pain over a long period. Complete compliance of recruitment in a primary care setting is often difficult to achieve in practice and one continuing concern is that there may have been selective recruitment of more severe cases. There was, in fact, wide variability in the numbers recruited between practices, consistent with variable compliance. We have previously shown[4] that there were no consistent differences in baseline severity between patients in practices who were high recruiters during the study, compared with practices with low recruiters. Similarly, as inevitably is the case with studies of this kind, recruitment declined during the period of study. Again we showed that there was no consistent difference in the pattern of severity between patients recruited early and those recruited late.[4] Finally, although the severity of disease in the patients recruited in this study may not be representative of all new attenders, the relationships observed between the predictors and being put on a waiting list are unlikely to differ.

A 'working definition' was used to define a new episode of hip pain. The localisation of hip as a source of pain is difficult in clinical practice and we relied on the use of pain drawings together with the GP's clinical judgement that the pain arose from structures within the hip joint. This definition was agreed upon by the GP participants at a consensus meeting prior to the study. In a pilot exercise involving 'paper patients' and 11 GPs, this proved to be reliable within

observers ($\kappa = 1.0$) as well as between observers ($\kappa = 0.6$). This was a study of hip pain attenders rather than of osteoarthritis itself. The results showed that radiographic damage, as would be entirely expected, was a key predictor in these patients being referred for surgery. Indeed, of 67 patients with a Croft radiographic score grade of three points or worse, 29 (43%) were put on a waiting list for joint replacement. In comparison, of the 128 patients with a Croft grade two or less, only three (2%) were put on a waiting list for replacement. However, interestingly, non-radiographic aspects of these patients were also important in explaining their future likelihood of being referred for surgery. In particular, restriction of internal rotation, as ascertained by the GP, did prove to be a very strong predictor. We have previously shown that this measurement is highly reproducible in general practice.[8]

The nature of the recruitment process meant that we were not able to truly identify the population incidence of new attendance, particularly as the participants were often individual GPs in multi-member partnerships. Such data would be necessary to extrapolate these findings to actual numbers of individuals in primary care and their rate of needing hip joint replacement.

Relationship to other work

Guidelines for assessing need for arthroplasty, such as those from the United States[11] and New Zealand,[12] have emphasised that pain and disability are the key factors that should be used. The US guidelines are difficult to put into operation and those from New Zealand, though based on a numerical score, are difficult to apply in practice.[13] In one hospital series from France of patients referred with osteoarthritis, severity of radiographic change was the major predictor of subsequent surgery.[14] The current study did not set out to investigate what factors influence primary care physicians to refer, and orthopaedic surgeons to accept, patients for arthroplasty. Thus we cannot judge as to what were the relative combinations of clinical and radiographic aspects, together with psychological, economic and co-morbidity features (though some of these were 'captured' within the different SF36 domains) in these decisions. The data do suggest, however, that the consequences in terms of patient numbers are similar for those who make decisions on clinical or on radiological grounds.

Implications of findings

Our main findings have linked these baseline clinical and radiographic features to one outcome of general practice care for patients who present with hip pain, namely arthroplasty. The benefits and usefulness of arthroplasty and the appropriateness of selection of patients for operation are not addressed by this study. The importance of the findings is that they provide epidemiological support for the observation that arthroplasties are currently being done in those patients who at their first visit have more severe disease. Whether severe disease is defined by X-rays, pain or disability, the outcome in terms of current replacements is similar. Future questions could address the issue of whether the baseline severity is related to the effectiveness of surgery, whether X-rays are necessary to define those who might benefit, and whether the current choice of patients for arthroplasty is the most effective in the context of all patients who present to their GP with pain in the hip. We can conclude, however, that it is likely, given the general agreement between surgeons and GPs when prioritising patients for joint replacement,[15] that characteristics at first consultation could be used to identify those who will need to be prioritised for surgical referral.

In summary, we have demonstrated in the primary care cohort we have studied, that hip joint replacement is quite frequent after first attendance with a new episode of hip pain and that radiographic findings and readily ascertainable clinical features predict which patients are likely to receive joint replacement in the short term.

References

1. Birrell F, Johnell O, Silman A. Projecting the need for hip replacement over the next three decades: influence of changing demography and threshold for surgery. *Ann Rheum Dis* 1999; **58:** 569-572.
2. Fox KM, Hochberg MC, Resnik CS, *et al.* Severity of radiographic findings in hip osteoarthritis associated with total hip arthroplasty. *J Rheumatol* 1996; **23(4):** 693-697.
3. Lawrence JS. *Rheumatism in populations*. London: William Heinemann Medical Books Ltd, 1977.
4. Birrell F, Croft P, Cooper C, *et al*. Radiographic change is common in new presenters in primary care with hip pain. PCR Hip Study Group. *Rheumatology* (Oxford) 2000; **39(7):** 772-775.
5. Birrell F, Croft P, Cooper C, *et al.* Health impact of pain in the hip region with and without radiographic evidence of osteoarthritis: a study of new attenders to primary care. The PCR Hip Study Group. *Ann Rheum Dis* 2000; **59(11):** 857-863.
6. Birrell F, Croft P, Cooper C, *et al*. Predicting radiographic hip osteoarthritis from range of movement. *Rheumatology* (Oxford) 2001; **40(5):** 506-512.
7. Ware JE, Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF36). I. Conceptual framework and item selection. *Med Care* 1992; **30(6):** 473-483.
8. Croft PR, Nahit ES, Macfarlane GJ, Silman AJ. Interobserver reliability in measuring flexion, internal rotation, and external rotation of the hip using a plurimeter. *Ann Rheum Dis* 1996; **55:** 320-323.
9. Croft P, Cooper C, Coggon D. Case definition of hip osteoarthritis in epidemiologic studies. *J Rheumatol* 1994; **21(4):** 591-592.
10. Lee KL, Harrell FE, Jr, Tolley HD, Rosati RA. A comparison of test statistics for assessing the effects of concomitant variables in survival analysis. *Biometrics* 1983; **39(2):** 341-350.
11. NIH consensus conference: Total hip replacement. NIH Consensus Development Panel on Total Hip Replacement. *JAMA* 1995; **273(24):** 1950-1956.
12. Hadorn DC, Holmes AC. The New Zealand priority criteria project. Part 1: Overview. *BMJ* 1997; **314(7074):** 131-134.
13. Frankel S, Eachus J, Pearson N, *et al*. Population requirement for primary hip-replacement surgery: a cross-sectional study. *Lancet* 1999; **353(9161):** 1304-1309.
14. (14) Vinciguerra C, Gueguen A, Revel M, *et al*. Predictors of the need for total hip replacement in patients with osteoarthritis of the hip. *Rev Rhum Engl Ed* 1995; **62(9):** 563-570.
15. Dent THS, Cunningham SL, Hook WE, *et al*. The MODEL project: a scoring system to manage demand for cataract and joint replacement surgery. *Br J Gen Pract* 2001; **51:** 917-919.

Acknowledgements

The co-operation of the patients and recruiting GPs from the Primary Care Rheumatology Society is much appreciated. The PCR Hip Study Group comprises Drs Martin Allen, Trevor Birnie, Susan Bradbury, David Child, A S Chopra, Tim Davies, Beryl Dennis, Elizabeth Dickson, John Dickson, Ian Louis Foster, Roger Gadsby, Iain Gilchrist, Charmian Goldwyn, Martin W Hall, Allan Harris, B M Heap, Stan G Hill, M J L Hopkins, Gill Hosie, David Jobling, Martin John, Fred Kavalier, P Lindsay, John Machen, Ian Robertson, James D A Robertson, Werner Rumfeld, R W Shaw, Ros Simpson, Lorna K R Smith, T G Stammers, Michael Stevenson, Anil Vaghmaria, Hywel Watkin, Andrew Watson, David Wheeler, Jane Wilcock and Michaela Young. Funding was provided by Arthritis Research Campaign (ARC) core funding. Dr Fraser Birrell was an ARC Clinical Epidemiology Training Fellow.