

Benchmarking: the five year outcome of rheumatoid arthritis assessed using a pain score, the Health Assessment Questionnaire, and the Short Form-36 (SF-36) in a community and a clinic based sample

N J Wiles, D G I Scott, E M Barrett, P Merry, E Arie, K Gaffney, A J Silman, D P M Symmons

Abstract

Background—Treatment, and therefore outcome, of rheumatoid arthritis (RA) will improve in the next few years. However, improvement in outcome can only be judged against the probability of certain outcomes with current conventional treatment.

Aim—To document the five year outcome of RA in the late 1990s.

Setting—Norfolk Arthritis Register (NOAR).

Design—Longitudinal observational cohort study.

Methods—318 patients with recent onset inflammatory polyarthritis recruited by NOAR in 1990–91 completed five years of follow up. Four groups were assessed: the whole cohort, all those referred to hospital, those who satisfied criteria for RA at baseline, and those referred to hospital who satisfied criteria for RA at baseline. Outcome was assessed with a visual analogue scale for pain, the Health Assessment Questionnaire (HAQ), and the Short Form-36 (SF-36).

Results—Of the RA hospital attenders, 50% had a visual analogue scale pain score of 5 cm or less and an HAQ score of 1.125 or less. SF-36 scores were reduced in all domains. Results are presented as cumulative percentages.

Conclusions—These results can be used for comparison and to set targets for improvement.

(*Ann Rheum Dis* 2001;60:956–961)

certain outcomes with current conventional treatment. This paper describes the five year outcome of RA, measured using a pain score and two validated and widely used self administered questionnaires on physical function and quality of life, in a community based sample recruited in 1990–91. The results are presented for the whole sample and for the subgroup who were referred to hospital.

Methods

The Norfolk Arthritis Register (NOAR) aims at recruiting all adults (aged 16 and over) who consult a general practitioner (GP) in the former Norwich Health Authority with swelling of two or more joints lasting for at least four weeks and with an onset since 1 January 1989. The protocol has been described in detail elsewhere.⁶ In summary, referred patients are assessed by a metrologist, usually within two weeks of notification. The metrologist conducts a structured interview, examines the joints and takes blood for rheumatoid factor estimation. Patients are reviewed annually. The 1987 ACR criteria for RA⁷ are applied at baseline and at each annual assessment. They are applied “cumulatively”—that is, if a patient ever satisfies a particular criterion that result is carried forward to subsequent assessments.

This “benchmarking” exercise focused on three important aspects of disease outcome at five years from registration: pain, physical function, and health related quality of life (HRQoL). Pain was measured with a 10 cm visual analogue scale (VAS). The pain score was recorded to the nearest centimetre (range 0–10). Physical function was measured by the Stanford Health Assessment Questionnaire (HAQ),⁸ modified for use by British patients.⁹ The HAQ measures functional ability in eight domains. It gives a score ranging from zero (no disability) to three (severe disability). HRQoL was assessed by the UK version of the Short Form-36 (SF-36).^{10 11} The SF-36 comprises 36 questions covering eight health domains: physical function, role limitations due to physical problems, bodily pain, general health, vitality, social functioning, role limitations due to emotional problems, and mental health. The score for each domain is transformed to a scale ranging from zero (poor health) to 100 (good health). As recommended by the developers of the SF-36,¹² missing values were imputed from

ARC Epidemiology Unit, University of Manchester Medical School, Oxford Road, Manchester M13 9PT, UK

N J Wiles
A J Silman
D P M Symmons

Department of Rheumatology, Norfolk and Norwich Hospital, Norfolk NR1 3SR, UK

D G I Scott
P Merry
E Arie
K Gaffney

Norfolk Arthritis Register, St Michael's Hospital, Aylsham, Norfolk NR11 6NA, UK
E M Barrett

Correspondence to: Professor Symmons
deborah.symmons@man.ac.uk

Accepted 27 March 2001

The treatment of rheumatoid arthritis (RA) is passing through exciting times. The past 12 months has seen the launch in the UK of a new class of non-steroidal anti-inflammatory drugs (COX-2 specific inhibitors),¹ a new second line agent (leflunomide),^{2 3} and a new category of drugs—biological agents designed to block the action of key cytokines.^{4 5} These drugs, used in conjunction with other established treatments, targeted at appropriate patients early in disease, should lead to a measurably improved outcome for people with RA. However, in order to establish in the future that outcome has improved, it is important to know “where we started from”—that is, the probability of

Table 1 Patient characteristics at baseline (n=433)

Age at symptom onset (years), median (IQR)	56 (42–68)
Female (No (%))	280 (64.7)
Symptom duration (months), median (IQR)	5.4 (2.8–11.3)
Number of swollen joints, median (IQR)	6 (1–14)
HAQ* score, median (IQR)	0.75 (0.25–1.50)
RF* positive (No (%))	139/375 (37.1)
Satisfied 1987 ACR* criteria for RA (No (%))	208 (48.0)

*HAQ = Health Assessment Questionnaire; RF = rheumatoid factor; ACR = American College of Rheumatology.

the average of the completed items if over half the items for that domain had been completed.

Outcome measures completed by patients were chosen in order to avoid the issue of interobserver variation in measures such as joint counts, and so to facilitate comparison with other patient cohorts.

PATIENT SAMPLE

The study group was derived from 482 patients notified to the NOAR in 1990 and 1991. Forty nine were subsequently excluded because they were given a diagnosis other than RA, psoriatic arthritis, viral arthritis, or undifferentiated inflammatory polyarthritis by a rheumatologist. The remaining 433 patients were followed up for five years, during which time 44 (10.2%) died, 47 (10.9%) declined further follow up, and 24 (5.5%) were lost to follow up. Thus 318/433 (73.4%) patients completed five years of follow up. Of these, 302 (95.0%) completed a VAS for pain, 317 (99.7%) completed an HAQ, and 308 (96.9%) completed all or part of the SF-36 at the fifth anniversary assessment. Fifty one (16.6%) SF-36 forms were incomplete, but in 20 cases the missing data could be imputed. Two hundred and seventy seven patients (87.1%) provided sufficient information to calculate a score for each of the eight domains of the SF-36.

Of the 318 patients who completed five years follow up, 238 (74.8%) had been referred to hospital during the follow up period. Just over half of the group satisfied the 1987 American College of Rheumatology (ACR) criteria for RA⁷ at the time of notification to NOAR and 237 (74.5%) satisfied the ACR criteria when applied cumulatively. One hundred and thirty

eight (86.3%) of the 160 patients who satisfied criteria for RA at baseline were referred to hospital. During the follow up period, 169/318 (53.1%) patients received second line treatment or steroids, or both. We looked at the outcome at five years for four groups: the whole cohort (n=318); all those referred to hospital (n=238); those who satisfied the ACR criteria for RA at baseline (n=160); and patients with RA referred to hospital (n=138).

STATISTICAL ANALYSIS

Data were analysed with the Statistical Package for the Social Sciences (SPSS)¹³ and Stata.¹⁴ As the distribution of the scores for three domains (role-physical, role-emotion, and mental health) were highly skewed, median values (and interquartile range (IQR)) are reported for all eight domains. NOAR data were compared with data from a large UK population sample (16 054 adults), the 1996 Health Survey for England.¹⁵

Results

Table 1 shows the baseline characteristics of the 433 patients with inflammatory polyarthritis who were recruited in 1990–91. The median age at onset was 56 years (IQR 42–68) and two thirds of the cohort were women.

Compared with those who completed five years follow up, those who died were older at the time of symptom onset (median age 70.0 years *v* 54.5 years) and more likely to be male (54.5% *v* 33.0%) (table 2). Patients who subsequently died also had a higher median HAQ (1.375; IQR 0.375–2.06) at registration than those who were followed up for five years (0.75; IQR 0.25–1.50). Although those who died were more likely to be seropositive, they were not more likely to satisfy the ACR criteria for RA at baseline. Those who withdrew from the study or were lost to follow up did not differ in age or sex from those who completed the follow up period; they did have milder disease as measured by the number of swollen joints and the proportion who satisfied the ACR criteria at the time of notification to NOAR (table 2).

Table 2 Comparison of the baseline characteristics of study completers with study non-completers

Variable	Patients who completed five years' follow up (n=318)	Withdrew or lost to follow up (n=71)	Difference (95% CI)*	Died during follow up (n=44)	Difference (95% CI)†
Age (years), median (IQR)	54.5 (41–65)	53.0 (39–68)	p=0.552‡	70.0 (63.5–78.5)	p<0.001‡
Female (No (%))	213 (67.0)	47 (66.2)	0.8 (–11.4 to 12.9)	20 (45.5)	21.5 (5.9 to 37.1)
Symptom onset (months), median (IQR)	5.5 (2.9–11.2)	5.2 (2.7–11.9)	p=0.796‡	4.5 (2.5–9.3)	p=0.526‡
Number of swollen joints, median (IQR)	7 (2–14)	4 (1–12)	p=0.116‡	6 (0–14)	p=0.305‡
HAQ, median (IQR)	0.75 (0.25–1.50)	0.75 (0.00–1.375)	p=0.143‡	1.375 (0.375–2.06)	p=0.006‡
RF positive (No (%))	98/277 (35.4)	18/59 (30.5)	4.9 (–8.2 to 17.9)	23/39 (59.0)	–23.6 (–7.2 to –40.0)
RA—list method (No (%))	160 (50.3)	28 (39.4)	10.9 (–1.7 to 23.5)	20 (45.5)	4.9 (–10.8 to 20.6)

*Differences between study completers and patients who withdrew or who were lost to follow up during the five years.

†Differences between study completers and patients who died during follow up.

‡Mann-Whitney U test.

Table 3 Characteristics of study completers at five years

Subgroup	Whole cohort (n=318)	Hospital attenders (n=238)	RA at baseline (n=160)	RA—hospital attenders (n=138)
Female (No (%))	213 (67.0)	160 (67.2)	108 (67.5)	94 (68.1)
Age at follow up (years), median (IQR)	60 (47–71)	59.5 (47–71)	61 (50–72.5)	60.5 (50–72)
Symptom duration at follow up (years), median (IQR)	5.5 (5.2–5.9)	5.5 (5.2–5.9)	5.5 (5.2–6.0)	5.6 (5.2–6.0)
Pain VAS, median (IQR)	4 (2–6)	4 (2–6)	4 (2–7)	5 (2–7)
HAQ, median (IQR)	0.75 (0.25–1.50)	0.875 (0.25–1.75)	1.06 (0.375–1.81)	1.125 (0.375–1.875)

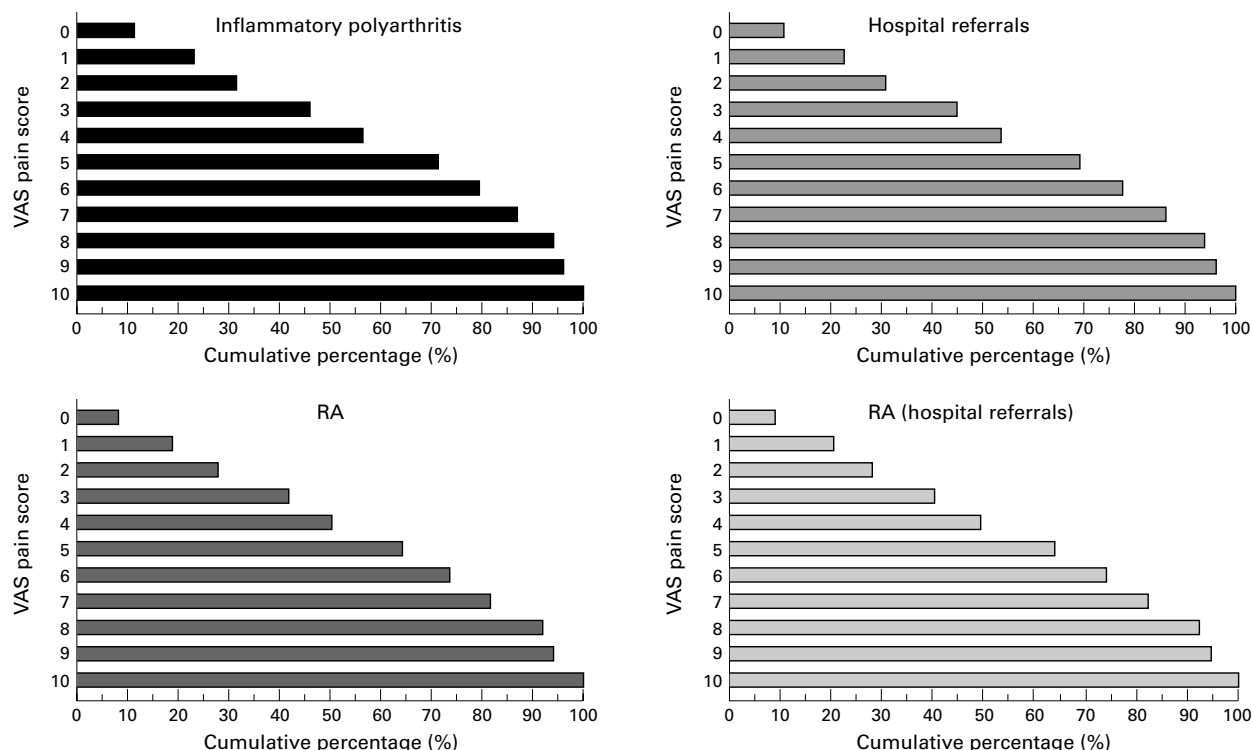


Figure 1 Visual analogue scale (VAS) pain score at five years.

The four groups analysed did not differ in their reported level of pain at five years (table 3). Figure 1 shows the cumulative proportion of each of the four patient groups who had a VAS pain score of a particular value or below. Thus, for example, 50% of hospital attenders with RA had a VAS pain score of 5 cm or more and 50% of all the other groups had a VAS pain score of 4 cm or more.

There was, however, a difference in HAQ scores between the groups (table 3). Figure 2 shows the cumulative proportion of each of the four patient groups who had an HAQ score of a particular value or below. Thus, for example,

30% of hospital attenders with RA had an HAQ score of 0.50 or less. Fifty per cent of the patients with inflammatory polyarthritis had an HAQ score of 0.75 or more at five years. In contrast, in the hospital attenders with RA, 50% had an HAQ of 1.125 or more. It is difficult to interpret precise values of the HAQ. However as a guide, an HAQ score of one or more is often used to represent “moderate” disability.^{16 17} Further details of the functional outcome and predictors of disability in this cohort have been published elsewhere.^{16 17}

All groups showed some impairment in HRQoL as measured by the SF-36 (table 4). The most seriously affected domain was the role-physical domain. The only difference between the groups was in the role-emotional domain. In the RA hospital attenders subgroup and the RA subgroup, 50% of patients had a score of zero in the role-physical domain. Although the score for each of the SF-36 domains ranges from 0 (poor health) to 100 (good health), the number of steps between the top and bottom of this range varies. For example, in the role-physical domain, the scoring system is such that it is only possible for the patient to score 0, 25, 50, 75, or 100. However, for other domains (for example, physical function and vitality) there are more than 20 possible scores. Figure 3 shows the results of the NOAR patients compared with those from a large sample of the UK population (controlled for age and sex).¹⁵ Compared with the general population, the NOAR patients (except for the RA hospital attenders) had little impairment in the mental health and role-emotional domains.

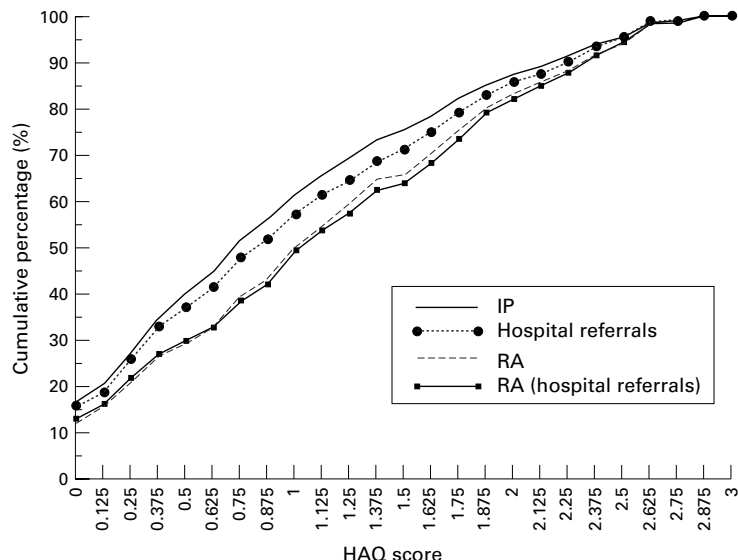


Figure 2 Health Assessment Questionnaire (HAQ) scores at five years. IP = inflammatory polyarthritis.

Table 4 SF-36 domain scores at the fifth anniversary assessment

Domain	Whole cohort		Hospital attenders		RA (at baseline)		RA—Hospital attenders	
	n	Median (IQR)	n	Median (IQR)	n	Median (IQR)	No	Median (IQR)
Physical function	298	55.0 (25–80)	221	50.0 (25–80)	148	45.0 (25–70)	126	45.0 (25–70)
Role—Physical	298	25.0 (0–100)	223	25.0 (0–100)	149	0.0 (0–100)	128	0.0 (0–100)
Bodily pain	300	55.6 (33.3–66.7)	225	44.4 (33.3–66.7)	147	44.4 (22.2–66.7)	127	44.4 (22.2–66.7)
General health	300	52.0 (35–72)	224	52.0 (32–72)	147	47.0 (32–72)	126	47.0 (32–72)
Vitality	302	45.0 (25–65)	227	45.0 (25–65)	151	40.0 (25–55)	130	40.0 (25–55)
Social functioning	302	75.0 (50–100)	226	75.0 (50–100)	149	75.0 (50–100)	128	75.0 (50–100)
Role—Emotional	300	100 (0–100)	224	100 (0–100)	148	83.3 (0–100)	127	66.7 (0–100)
Mental health	303	72.0 (60–84)	228	72.0 (56–84)	152	72.0 (58–84)	131	72.0 (56–84)

Discussion

Theoretically the patients presented in this report should include all new cases of inflammatory polyarthritis (and RA) who presented to primary care in the Norwich Health Authority in 1990–91. There will of course be underascertainment—mainly due to failure by GPs to notify the NOAR of patients who were not subsequently referred to hospital. We do not know the degree of underascertainment, though we are currently conducting a population based prevalence survey of RA in this area which should provide some insight into the number of missing cases. Some cases with an onset in 1990–91 will have presented to primary care

after 31 December 1991 and so this cohort is an inception cohort of early inflammatory polyarthritis that presented in 1990–91.

There have been relatively few reports of the five year outcome of RA using the HAQ^{18 19 20} and none using either a VAS pain score or the SF-36. Kvien *et al* used a VAS pain score in a survey of 1552 subjects on a community register of all hospital diagnosed cases of RA in Oslo.²¹ The mean VAS in the 1030 respondents (who had a mean disease duration of 13 years) was 4.6 cm (SD 0.08). This is similar to the median VAS of 5 cm found in our study, suggesting that pain score may vary little with disease duration. In another inception cohort

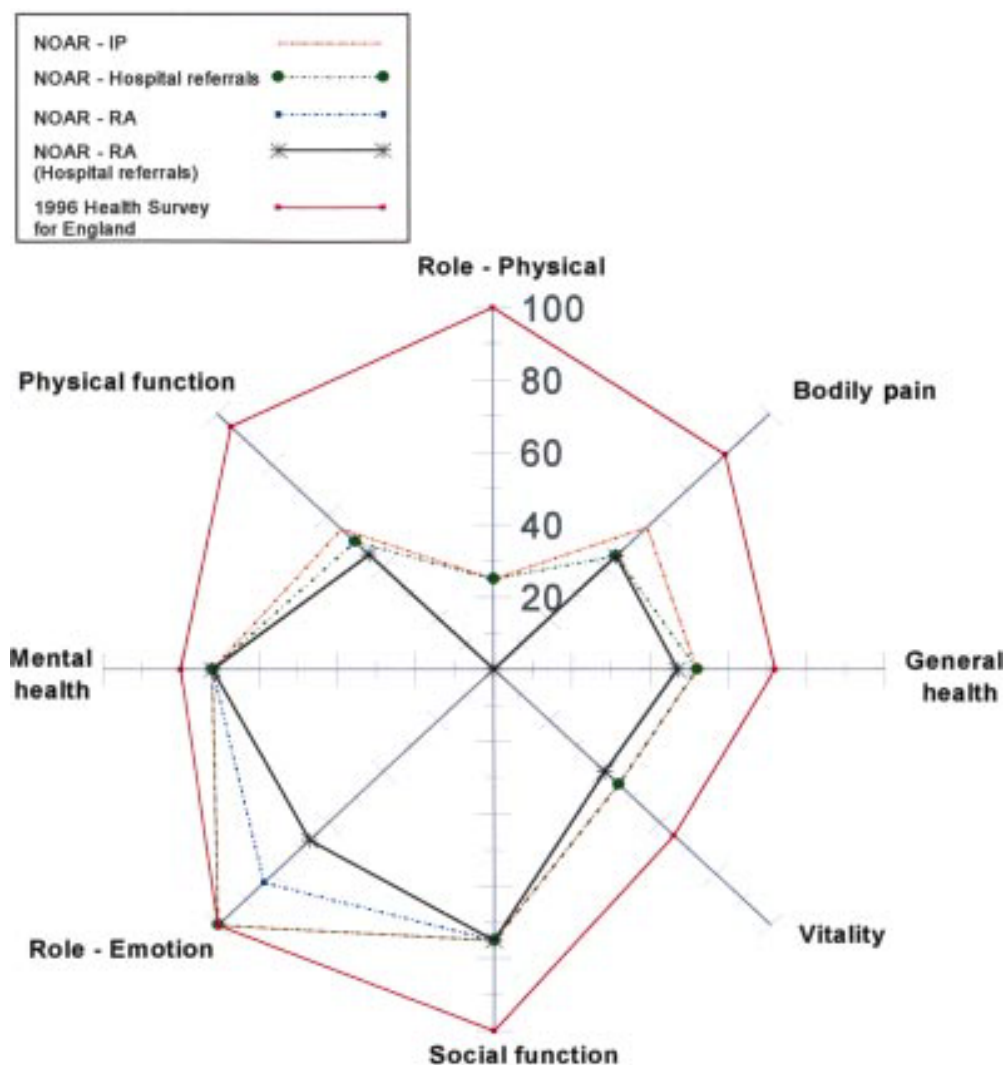


Figure 3 Comparison of NOAR SF-36 domain scores with UK population normative data. IP = inflammatory polyarthritis.

Table 5 Comparison of SF-36 data from NOAR with other studies of UK hospital patients with RA

Study (Reference)	NOAR-RA hospital attenders	Talamo et al ²²	Ruta et al ²³
Patients			
No	138	137	233
Mean age (range)	60 (24–87)	62 (26–91)	56 (21–87)
Mean disease duration, years (range)	5.7 (4.0–7.8)	11 (0–43)	13 (0.2–65)
Female (%)	68	71	81
Source of patients	Hospital referrals	Outpatient clinic	Outpatient clinic
Definition of RA	1987 Criteria*	1987 Criteria	1987 Criteria
Median domain scores			
Physical function	45.0	35.0	25.0
Role - Physical	0.0	0.0	0.0
Bodily pain	44.4	33.3	37.0
General health	47.0	47.0	41.0
Vitality	40.0	45.0	40.0
Social functioning	75.0	66.7	63.0
Role - Emotion	66.7	66.7	67.0
Mental health	72.0	68.0	72.0

*List criteria for RA satisfied at baseline.

Table 6 Comparison of SF-36 data from NOAR with other community based studies of RA

Study (references)	NOAR-RA Community	Norway ²¹	Australia ²⁴
Patients			
No	160	1030	122 (approx)
Mean age (range)	60 (24–88)	62 (20–79)	NA
Mean disease duration, years (range)	5.7 (4.0–7.8)	12.9 (SD 11.4)	NA
Female (%)	68	79	61
Source of patients	Primary care	Community register of hospital diagnosed RA	Population survey
Definition of RA	1987 Criteria		Self reported doctor diagnosis
Mean domain scores			
Physical function	47.2	47.3	72.7
Role - Physical	36.6	27.3	65.4
Bodily pain	49.4	41.0	58.6
General health	50.6	42.0	58.6
Vitality	42.0	39.4	52.7
Social functioning	70.0	63.7	78.5
Role - Emotion	59.7	52.0	83.4
Mental health	70.1	68.1	76.3

NA = not available.

study from Norway the mean HAQ at five years was 0.91 (SD 0.65) and 40% of patients had an HAQ score >1.²⁰ The median HAQ score at five years in patients recruited at each of the nine UK centres contributing to the Early Rheumatoid Arthritis Study (ERAS) ranged from 0.8 to 1.3.¹⁹ In seven of the nine centres the median HAQ was either 1.0 or 1.1. The median HAQ for the NOAR RA hospital attenders was 1.125.

Although there have been a number of publications (mainly clinical trials) which have used the SF-36 in patients with RA, few have reported the actual scores for each domain. Two of these studies involved UK hospital attenders (table 5). They represent the whole spectrum of disease duration, with a substantially higher median duration than the NOAR patients. This is reflected in the poorer median scores for physical function and bodily pain compared with the NOAR RA hospital scores. The patients on the Oslo community register of patients with hospital diagnosed RA, who also had a greater disease duration than the NOAR community patients with RA, generally had poorer scores. A population survey in Australia which asked respondents whether they had “ever been diagnosed as having RA by a doctor” found better scores in all domains of the SF-36 than any other published study of RA²⁴ (table 6). This raises the possibility of diagnostic misclassification.

We believe that the rate of referral and the management of hospital referred cases was “typical” for the UK at the time (1990–91). During the follow up period 169 (53%) patients received second line drugs or steroids. Sulfasalazine was the drug of first choice, though methotrexate became more popular during the follow up period. These data may, therefore, be regarded as typical of the five year outcome of RA and inflammatory polyarthritis in the UK in the late 1990s. Figures 1 and 2 could be used to set targets for better treatment in the future—and this in turn offers possibilities for RA to be included in National Programmes for Health Improvement. For example, from fig 2, it can be seen that 40% of hospital attenders with RA have an HAQ score <0.75 at five years. It would then be possible to set a target that 60% of such patients should have an HAQ score <0.75 at some future date. Similarly, this could provide a standard against which other departments could assess the outcome of their patients. Any differences in outcome between centres would then offer avenues for further research.

- Emery P, Zeidler H, Kvien TK, Guslandi M, Naudin R, Stead H, et al. Celecoxib versus diclofenac in long-term management of rheumatoid arthritis in randomised double-blind comparison. *Lancet* 1999;354:2106–11.
- Smolen JS, Kalden JR, Scott DL, Rozman B, Kvien TK, Larsen A, et al. Efficacy and safety of leflunomide compared with placebo and sulphasalazine in active rheumatoid arthritis: a double-blind randomised multicentre trial. *Lancet* 1999;353:259–66.

- 3 Emery P, Breedveld FC, Lemmel EM, Kaltwasser JP, Dawes PT, Gornor B, *et al.* A comparison of the efficacy and safety of leflunomide and methotrexate for the treatment of rheumatoid arthritis. *Rheumatology (Oxford)* 2000;39:655–65.
- 4 Moreland LW, Schiff MH, Baumgartner SW, Tindall EA, Fleischmann RM, Bulpitt KJ, *et al.* Etanercept therapy in rheumatoid arthritis: a randomised, controlled trial. *Ann Intern Med* 1999;130:478–86.
- 5 Maini RN, St Clair EW, Breedveld F, Furst D, Kalden J, Weisman M, *et al.* Infliximab (chimeric anti-tumour necrosis factor alpha monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomised phase III trial. *Lancet* 1999;354:1932–9.
- 6 Symmons DPM, Barrett EM, Bankhead CR, Scott DGI, Silman AJ. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. *Br J Rheumatol* 1994;33:735–9.
- 7 Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, *et al.* The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
- 8 Fries JF, Spitz P, Kraines RE, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23:137–45.
- 9 Kirwan JR, Reeback JS. Stanford Health Assessment Questionnaire modified to assess disability in British patients with rheumatoid arthritis. *Br J Rheumatol* 1986;25:206–9.
- 10 Jenkinson C, Layte R, Wright L, Coulter A. *The UK SF-36. An analysis and interpretation manual.* Oxford: Joshua-Horgan Print Partnership, 1996.
- 11 Brazier JE, Harper R, Jones NMB, O’Cathain A, Thomas KJ, Usherwood T, *et al.* Validating the SF-36 health survey questionnaire: new outcome measures for primary care. *BMJ* 1992;305:160–4.
- 12 Ware JEJ, Sherbourne CD. The MOS 36-item short form health survey (SF-36). *Med Care* 1992;30:473–83.
- 13 SPSS Inc. *Statistical package for the social sciences.* Version 6.1. Chicago, 1991.
- 14 StataCorp. *Stata statistical software.* Release 5.0. College Station, TX: Stata Corporation, 1997.
- 15 Prescott-Clarke P, Primatesta P, eds. *Health survey for England 1996.* London: The Stationery Office, 1998.
- 16 Wiles NJ, Dunn G, Barrett EM, Harrison BJ, Silman AJ, Symmons DPM. One year follow-up variables predict disability 5 years after presentation with inflammatory polyarthritis with greater accuracy than at baseline. *J Rheumatol* 2000;27:2360–6.
- 17 Wiles N, Dunn G, Barrett E, Silman A, Symmons D. Associations between demographic and disease-related variables and disability over the first five years of inflammatory polyarthritis: a longitudinal analysis using generalised estimating equations. *J Clin Epidemiol* 2000;53:988–96.
- 18 Eberhardt KB, Fex E. Functional impairment and disability in early rheumatoid arthritis - development over 5 years. *J Rheumatol* 1995;22:1037–42.
- 19 Young A, Dixey J, Cox N, Davies P, Devlin J, Emery P, *et al.* How does functional disability in early rheumatoid arthritis affect patients and their lives? Results of 5 years of follow-up in 732 patients from the early RA study. *Rheumatology (Oxford)* 2000;39:603–11.
- 20 Uhlig T, Smedstad LM, Vaglum P, Mourm T, Gérard M, Kvien TK. The course of rheumatoid arthritis and predictors of psychological, physical and radiographic outcome after 5 years of follow-up. *Rheumatology (Oxford)* 2000;39:732–41.
- 21 Kvien TK, Kaasa S, Smedstad LM. Performance of the Norwegian SF-36 Health Survey in patients with rheumatoid arthritis. II. A comparison of the SF-36 with disease-specific measures. *J Clin Epidemiol* 1998;51:1077–86.
- 22 Talamo J, Frater A, Gallivan S, Young A. Use of the short-form 36 (SF36) for health status measurements in rheumatoid arthritis. *Br J Rheumatol* 1997;36:463–9.
- 23 Ruta DA, Hursat MP, Kind P, Hunter M, Shillings A. Measuring health status in British patients with rheumatoid arthritis: reliability, validity and responsiveness of the Short Form 36-Item Health Survey (SF-36). *Br J Rheumatol* 1998;37:425–36.
- 24 Hill CL, Parsons J, Taylor A, Leach G. Health related quality of life in a population sample with arthritis. *J Rheumatol* 1999;26:2029–35.