

EXTENDED REPORT

Comparing the long-term clinical outcome of treatment with methotrexate or sulfasalazine prescribed as the first disease-modifying antirheumatic drug in patients with inflammatory polyarthritis

S L Hider, A Silman, D Bunn, S Manning, D Symmons, M Lunt



Ann Rheum Dis 2006;65:1449–1455. doi: 10.1136/ard.2005.049775

Objective: To compare the clinical and functional outcome at 2 and 5 years in patients with inflammatory polyarthritis treated with either methotrexate (MTX) or sulfasalazine (SSZ) as the first disease-modifying antirheumatic drug (DMARD).

Methods: Patients recruited to a primary-care-based inception cohort of patients with inflammatory polyarthritis were eligible for this analysis if they were started on either SSZ (n = 331) or MTX (n = 108) as their first DMARD within 3 months. Outcomes assessed included the Disease Activity Score (DAS)28, Health Assessment Questionnaire, radiological erosions (Larsen Score) and cumulative mortality with the proportions still on the original treatment. To overcome potential bias in allocation to these two treatments, a propensity score was calculated based on baseline disease status variables. Results are expressed as the mean difference between MTX and SSZ, both unadjusted and adjusted for propensity score.

Results: The baseline differences between the two groups disappeared after adjusting for propensity score. At 2 and 5 years there were few differences in the clinical outcomes, either unadjusted or after adjustment for propensity. By contrast, at 5 years the proportion that was erosive was lower in the MTX group: odds ratio 0.3 (95% confidence interval 0.1 to 0.8), with a 31% lower Larsen Score after adjustment. At both time points, those treated with MTX were at least twice as likely to remain on that drug as those treated with SSZ.

Conclusion: Long-term clinical outcome is similar in patients prescribed MTX and SSZ, although it would seem that MTX has greater potential to suppress erosions, which supports it being the first DMARD of choice.

See end of article for authors' affiliations

Correspondence to:
M Lunt, Arc Epidemiology Unit, Stopford Building, University of Manchester, Oxford Road, Manchester M13 9PT, UK;
mark.lunt@manchester.ac.uk

Accepted 2 March 2006
Published Online First
15 March 2006

It is now well-accepted that early and aggressive use of disease-modifying antirheumatic drugs (DMARDs) can improve the long-term clinical outcome¹ in patients with rheumatoid arthritis. However, it is less clear whether there is an optimum therapeutic strategy of combination therapy versus monotherapy or, indeed, which drug should be prescribed first. The choice of initial treatment is even more critical as recent studies suggest that the first DMARD given is likely to be the most effective in terms of reducing C reactive protein (CRP) levels.²

Many factors, such as relative efficacy, ease of administration, side-effect profile, monitoring requirements and cost, influence the choice of a drug. In addition, assessment of prognosis means that patients with very active disease tend to be given the most aggressive treatment. Surveys suggest that although methotrexate (MTX) is the most commonly preferred first drug,² sulfasalazine (SSZ) remains the first choice among some rheumatologists.³ A meta-analysis of clinical trial data suggested that both agents had similar efficacy.⁴ Recent clinical trials have provided an opportunity to study outcome from both of these agents when compared with newer treatments such as anti-tumour necrosis factor (TNF) treatments or leflunomide. These studies suggest that typical American College of Rheumatology 20 responses for SSZ and MTX are similar: range 44–59% and 46–65%, respectively,^{5–8} when comparing data between clinical trials.

However, to date only two studies have directly compared the clinical effectiveness of SSZ and MTX in patients with

early rheumatoid arthritis.^{6–9} Both of these were 1-year studies, in patients who were DMARD-naïve and rheumatoid factor-positive, comparing treatment with a combination of MTX and SSZ with treatment with the individual drugs. No significant differences were seen at 1 year in Disease Activity Score (DAS)28 or in the proportion satisfying the American College of Rheumatology or European League Against Rheumatism criteria for response. A subsequent report followed up 146 of the 205 patients included in one of the original trials⁶ and compared the 5-year outcome in patients randomised to either combination therapy or monotherapy, either SSZ or MTX.¹⁰ No significant difference was observed between the SSZ and MTX group, although the full results were not reported as the primary aim was to make a comparison between combination therapy and monotherapy rather than between the individual monotherapies.

Thus, there are few data comparing the long-term outcome of the use of these two agents as the preferred DMARD. Clearly, it is both difficult and expensive to maintain long-term follow-up in the constraints of a randomised clinical trial and so longer-term outcomes are increasingly being studied using prospective observational cohorts. As these are non-randomised, bias in allocation to one treatment over

Abbreviations: CRP, C reactive protein; DAS, Disease Activity Score; DMARD, disease-modifying antirheumatic drug; HAQ, Health Assessment Questionnaire; MTX, methotrexate; NOAR, Norfolk Arthritis Register; SJC, swollen joint count; SSZ, sulfasalazine; TJC, tender joint count; TNF, tumour necrosis factor

another may occur (known as “confounding by indication”). The propensity score has been proposed as a method for adjusting for potential bias in allocation of treatment¹¹ with outcomes compared between treatment groups after adjusting for propensity score.

Our study aimed to investigate whether there were differences in the 2-year and 5-year outcome in a large inception cohort of patients prescribed MTX or SSZ as their first DMARD.

PATIENTS AND METHODS

Study design

A prospective cohort study was undertaken following up patients recruited to a large population-based registry of patients with inflammatory polyarthritis recruited from primary care. Patients treated with either SSZ or MTX as their first DMARD were followed up at 2 and 5 years to compare their clinical, functional and radiological outcome.

Patients

The patients for this study were recruited from the Norfolk Arthritis Register (NOAR), a primary-care-based inception cohort of patients with early inflammatory polyarthritis. Details of the registry have been published elsewhere.¹² Briefly, the registry covers the area of the former Norfolk Health Authority, with a population of roughly half a million people. Any adult (age ≥ 16 years) presenting to his or her primary care physician with swelling of two or more joints lasting at least 4 weeks is notified to NOAR. A parallel notification system operates from hospitals within the catchment area. All patients who are referred to secondary care are treated by one of three rheumatologists according to their routine clinical practices.

Between 1990 and 1999, 2659 patients were recruited by NOAR. NOAR aims to undertake the baseline assessment within 2 weeks of receiving notification. At the hospital, attending patients were referred to NOAR after their first hospital attendance; some patients would have been started on DMARD treatment before notification. For the purposes of this analysis, investigation is restricted to patients who started treatment with either SSZ or MTX as their first DMARD within 3 months of their baseline visit and had been followed up for at least 2 years. Patients started on MTX at 7.5 mg/week or on SSZ at 2 g/day. Details of the patient cohorts and their follow-up are summarised in fig 1A,B. All patients provided informed consent. The local research ethics committee approved the study.

Assessment

One of a team of trained research nurses carried out a structured interview and clinical examination at the baseline and follow-up assessments. Data recorded included age at symptom onset, sex, symptom duration, tender joint count (TJC, 53 joints) and swollen joint count (SJC, 51 joints). At baseline, blood samples were taken for rheumatoid factor and CRP measurement. Rheumatoid factor was measured using a latex agglutination technique, and a titre of $\geq 1/40$ was considered positive. The DAS28 was calculated for each patient using the 28 SJC, 28 TJC and CRP, using the following formula¹³:

$$\text{DAS28} = (0.56 \times \text{SQRT}(\text{TJC}28) + 0.28 \times \text{SQRT}(\text{SJC}28) + 0.36 \times \ln(\text{CRP}+1)) \times 1.10 + 1.15.$$

All patients also completed a Health Assessment Questionnaire (HAQ), modified for use in British patients.¹⁴ At 5 years, patients had a repeat CRP and the DAS28 was calculated. At this point, patients were also invited to have x rays of the hands and feet, which were scored for the

presence or absence of erosions (and patients were classified as having erosions or not) and the Larsen Score as previously described.¹⁵ At both follow-up points, the proportion of patients who were still on their original treatment was derived as an indicator of the combined effects of tolerance and efficacy. For the purposes of this analysis, remission was defined as no swollen or tender joints in patients not currently taking DMARDs.

Statistical analysis

The baseline demographics, disease activity and severity measures were compared between the two groups. The outcomes were expressed as changes from baseline in the joint counts, CRP, DAS28 and HAQ score. Differences in remission, still on treatment and erosive (at 5 years) were expressed as odds ratios (ORs) (with 95% confidence interval (CI)). The Larsen Scores were compared at 5 years.

As this was an observational study there was likely to be confounding by indication in allocation between the two treatments. We therefore adjusted for such differences using a propensity score, which predicts the probability of receiving MTX rather than SSZ as a first treatment, from all baseline variables that might also predict the outcome. Patients were then divided into strata, on the basis of their propensity score. Within each stratum, the baseline variables between the two treatments should be balanced. Thus, any within-stratum difference in outcome is assumed to be due to the differences in treatment.

Normally, logistic regression is used to calculate the propensity score. However, if the model is not specified correctly, the propensity score will not balance the baseline variables. It was therefore essential to check the fit of the propensity score model. The Hosmer–Lemeshow test¹⁶ was used to verify that the differences between the observed and expected numbers of patients receiving each treatment within each stratum were small enough to be explained by chance alone, and would suggest that no variables that differed greatly between the two treatment groups had been omitted from the model. Secondly, the balance of each baseline variable between the two treatments within each stratum was assessed: if any variable remains unbalanced, it suggests that there is an interaction between that variable and another that should be entered into the logistic regression model, or that the association between that variable and treatment assignment is non-linear, and a more complex model needs to be fitted. For continuous variables, the balance before and after stratifying by propensity score was calculated using Wilcoxon’s rank sum test and van Elteren’s test (the respective non-parametric equivalents of one-way and two-way analyses of variance). For categorical variables, logistic regression was used.

Comparing outcomes

The outcomes at 2 and 5 years were compared using Wilcoxon’s rank sum test for continuous variables and the χ^2 test for categorical variables. As a more powerful test, the outcomes in the two treatment groups were then compared after adjustment for the relevant baseline values (except for the radiographic outcome, as x rays were not taken at baseline). Linear regression was used for all continuous variables to compare the mean change and logistic regression used for the categorical variables, with results expressed as ORs. Finally, these regression analyses were repeated after adjusting for propensity stratum. The residuals for all models of continuous variables approximated to a normal distribution, except for the Larsen Score. This variable was highly skewed, and thus the log score was used as the dependent (outcome) variable.

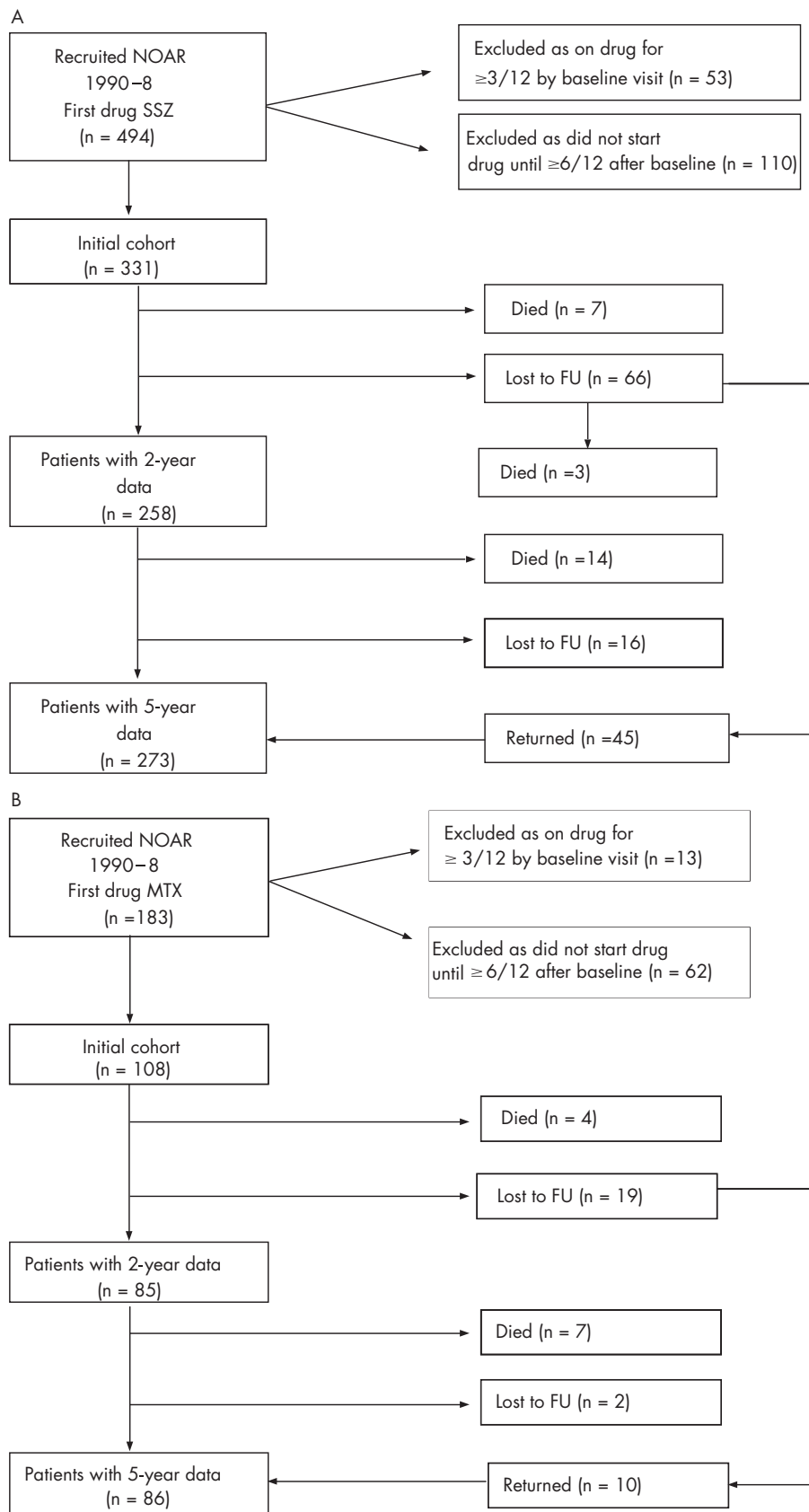


Figure 1 Flow chart of participants using (A) sulfasalazine (SSZ) and (B) methotrexate (MTX). FU, follow-up; NOAR, Norfolk Arthritis Register.

RESULTS

Characteristics of the study cohort at baseline

In all, there were 677 patients whose first treatment was either MTX or SSZ; of these, 66 had been on treatment for >3 months by the time of their baseline visit and 172 did not start treatment until >6 months after baseline, leaving 439 for subsequent analysis. The numbers of patients on each drug who were successfully followed up at the two time points are shown in figs 1A,B. Strenuous efforts were made at 5 years to contact those patients lost to follow-up at 2 years, and so the follow-up rates were higher at the later follow-up. Data were available on 78% and 79% of all eligible patients in the SSZ and MTX groups, respectively.

Table 1 shows the baseline characteristics of each group. Patients starting SSZ were younger at diagnosis (median age 53 v 58 years; $p = 0.1$) and had higher SJC and TJC (medians 8 v 7; $p = 0.06$ and 10 v 7; $p = 0.1$, respectively). The proportion who were coprescribed steroids was higher in the MTX group (14, 13.9%) than in the SSZ group (22, 6.7%). Baseline CRP, DAS28 and HAQ scores were similar between the two groups. The most likely explanation for these differences lies in a secular change with both an increasing use of MTX and a lower threshold, in terms of the number of active joints, for starting either agent as the 10-year recruitment period progressed (fig 2).

Propensity scoring

The propensity score was calculated using only those 358 patients who were followed up for 5 years. In addition to the variables listed in table 1, for the reasons stated earlier (fig 2), the date at which the patient was registered by NOAR was also included in the propensity score. As the increase in the use of MTX was non-linear with time, a quadratic term for time was also included in the model. Apart from time, the only other significant predictor of MTX treatment was age at onset, with older patients being more likely to be treated with MTX. Patients satisfying rheumatoid arthritis criteria at baseline were more likely to receive treatment with MTX, but the effect was not significant (OR 1.9, 95% CI 0.96 to 3.8).

Figure 3 shows the proportion of patients treated with MTX in each of the propensity score quintiles, increasing from 1% in the bottom quintile to 72% in the highest. The Hosmer-Lemeshow χ^2 was 1.2 (df = 3, $p = 0.7$), which suggests that the model fitted the data extremely well. As a further test, all variables in the propensity score were balanced between treatments, having adjusted for propensity quintile (table 1).

Outcomes after 2 years

Table 2 shows the outcome at 2 years. In both the groups, there were declines in both SJC and TJC, and in mean HAQ, although the changes were very similar with the two drugs. A higher proportion of patients taking MTX (56%) had no change in treatment over the 2 years compared with those taking SSZ (50%), which after adjustment was equivalent to an approximate doubling of the odds of drug survival in the MTX group. Adjusting for propensity (table 2) made little difference to the results. There was, for example, a propensity-adjusted difference of 0.03 lower fall in HAQ in the MTX than in the SSZ groups.

Outcomes after 5 years

At 5 years, in addition to the clinical data described earlier, information was also available on CRP (enabling calculation of the DAS28 score) and radiological erosions. There were larger falls in both the TJC and SJC in the SSZ-treated group than in the MTX-treated group, which were significant (table 3). These joint counts were higher at baseline in the SSZ-treated group and, after adjusting for the propensity

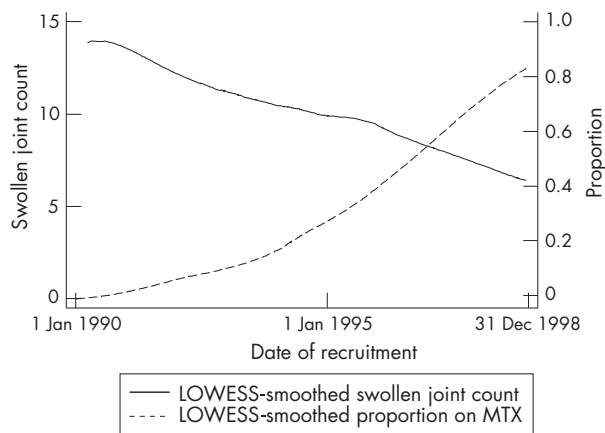


Figure 2 Changes in mean swollen joint count and proportion of patients treated with methotrexate (MTX) over time. LOWESS is a data analysis technique for producing a “smooth” set of values from a time series that has been contaminated with noise, or from a scatter plot with a “noisy” relationship between the two variables. In a time series context, the technique is an improvement over least-squares smoothing when the data are not equally spaced (as least-squares smoothing assumes; http://www.itl.nist.gov/div898/software/dataplot/refman1/ch3/lowess_s.pdf).

score and the baseline score, there was no significant difference in the change in either count between the treatments (table 3). The mean CRP showed a greater increase in the MTX group, as did the mean 5-year DAS28, although this difference was not significant, and after adjustment the changes in 5-year DAS28 were almost identical (table 3).

The proportions of erosive and mean Larsen Scores were higher in the SSZ group, although these crude differences were not significant. The odds of erosions at 5 years in the MTX group were 0.8 of those of the SSZ group before adjusting for propensity, but this apparent protection increased to 0.3 (95% CI 0.1 to 0.8) after adjustment (table 3). The mean difference in Larsen Score also increased after adjustment and was 31% lower in the MTX group, although this was not significant.

More patients in the MTX group were still taking MTX as their first DMARD (34% v 22%), as had been the case at 2 years. After adjusting for propensity this was equivalent to a 2.2-fold greater odds of drug survival (table 3).

In general, adjusting for the propensity score tended to improve the outcome of the patients treated with MTX, suggesting, given the baseline disease status, that the predicted (expected) outcome of these patients was worse than the expected outcome in the patients treated with SSZ.

DISCUSSION

This study recruited 439 patients with inflammatory polyarthritis as soon as possible after their first attendance at primary care, who were followed up for 5 years, having started treatment, within 6 months of symptom onset, with either MTX or SSZ as their first DMARD. Approximately 75% of these had been treated with SSZ. Clear secular patterns were observed during the 10-year recruitment period, with a shift towards greater first use of MTX and towards starting DMARDs with fewer active joints. We found few other differences in the disease characteristics between those prescribed these two agents. Continuation on MTX was substantially more in patients treated with MTX and this agent was also associated with a considerable reduction in the likelihood of erosions by 5 years. Baseline radiographic data were not collected in these patients and therefore we

Table 1 Baseline characteristics of each group

	MTX group, n = 108	SSZ group, n = 331	p Value for difference	
			Unadjusted for propensity	Adjusted for propensity
Age at diagnosis (years), median (IQR)	58 (39, 67)	53 (41, 63)	0.09	0.7
Female, n (%)	69 (64)	200 (60)	0.6	0.9
RF positive, n (%)	48 (47)	137 (46)	0.8	0.6
Satisfy RA criteria at baseline, n (%)	60 (57)	176 (60)	0.9	0.9
SJC (51 joints), median (IQR)	7 (2, 14)	8 (3, 15)	0.06	0.9
TJC (53 joints), median (IQR)	7 (2, 17)	10 (4, 20)	0.1	0.8
Nodules, n (%)	11 (10)	28 (9)	0.9	0.6
Baseline CRP*, median (IQR)	1.4 (3, 30)	1.2 (5, 30)	0.8	0.5
Baseline DAS28, median (IQR)	4.5 (3.0, 5.4)	4.6 (3.6, 5.5)	0.3	0.8
Baseline HAQ, median (IQR)	1.3 (0.5, 1.9)	1.1 (0.5, 1.6)	0.6	0.9
Delay from symptom onset to drug start (months), median (IQR)	5.8 (3.0, 13.1)	7.1 (3.1, 11.2)	0.9	0.9

CRP, C reactive protein; DAS, Disease Activity Score; HAQ, Health Assessment Questionnaire; IQR, interquartile range; MTX, methotrexate; RA, rheumatoid arthritis; RF, rheumatoid factor; SJC, swollen joint count; SSZ, sulfasalazine; TJC, tender joint count.
*93 of MTX group and 261 of SSZ group had CRP measured at baseline.

were unable to look at change by, as opposed to state at, 5 years in radiographic appearance. Thus the differences observed may represent baseline effects rather than drug effects. It was assumed that the likelihood of radiographic damage, given the short duration and primary care recruitment, would be small at baseline, which, if correct, would suggest that the adjusted 5-year differences were real. Further, the baseline disease characteristics, such as the median DAS and rheumatoid factor positivity, were virtually identical in the two treatment groups (table 1) and hence the expected erosion risk at baseline was unlikely to have been considerably different.

This was not a randomised clinical trial but an observational study of treatment outcome. It was also a study of the effect of the first DMARD prescribed and, as shown above, considerable proportions of patients at both the 2-year and 5-year points were not taking this first agent: some had ceased DMARD treatment and others had switched. This analysis can be considered to have dealt with the question “do patients in the ‘real world’ who start with MTX as their first DMARD have a better outcome than those who start with SSZ, independent of what treatment changes happen subsequently?”

As this was not a clinical trial, the doctor was free to change, stop or add treatments as clinically indicated. The starting doses of 7.5 mg/week for MTX and 2 g/day for SSZ would have changed during the course of the study in

relation to perceived change in clinical state. Further, as may be expected, patients changed their actual drug during the follow-up. Thus, 41% of patients starting on SSZ had switched to MTX at some stage, with only 13% changing in the opposite direction. An observational study such as this cannot easily adjust for the confounding by indication that leads to treatment change during the course of observation. Indeed, such changes in response to changed clinical state should result in greater outcome similarity and, for example, would be against the direction of finding a beneficial effect on radiological outcome from starting with low-dose MTX.

In NOAR the assessments are made by specially recruited nurses but the therapeutic decisions are made by the patient’s own rheumatologist. Thus, the baseline disease status used in this paper does not reflect the status on the actual date of starting treatment. We allowed up to 3 months between baseline and starting treatment to keep the delay as short as possible. As implied earlier, there were several patients who were started on DMARD at their first visit and hence were already on treatment by their baseline visit. Although this would have had a major effect, such an effect, if present, would be greater earlier on. We therefore undertook a subgroup analysis on outcome at 2 years, excluding those patients who were referred to NOAR after they had also been started on treatment. We found no major differences between the subgroup and the dataset as a whole. Thus, the proportions still on treatment for the whole cohort and after excluding those already on treatment were 56% and 54% for MTX and 46% and 49% for SSZ. Changes in the other parameters measured (table 2) were also virtually identical.

The comparison would also have been influenced if there were considerable differences in baseline cotherapy with other DMARDs, including steroids, between the groups. The proportions taking another DMARD at this stage were small and the other agent was almost always a steroid. The proportion taking steroids was higher in the MTX group (14, 13.9%) than in the SSZ group (22, 6.7%). Adjusting for these baseline treatments, in addition to the propensity scores, did not alter the results in any substantive way (data not shown).

Apart from the secular trend towards the use of MTX over SSZ, there were only a few differences in the baseline disease characteristics, with some evidence towards a policy of starting DMARD treatment with fewer active joints.

The approach used in this study had some advantages. We were able to follow up patients for 5 years, an objective not easily achieved in a clinical trial. The population was unselected as far as possible and the results should have wide applicability to other newly recognised cohorts with inflammatory polyarthritis. We have shown elsewhere that

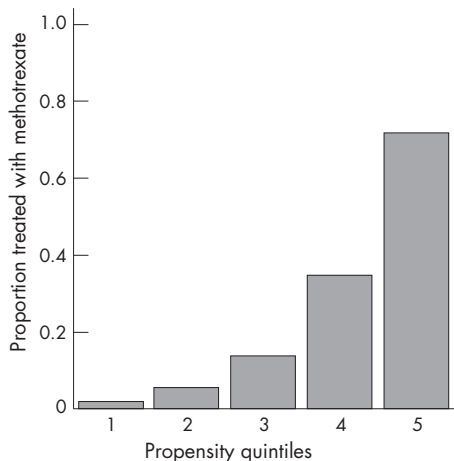


Figure 3 Proportion of patients treated with methotrexate in each of the propensity quintiles.

Table 2 Two-year outcome for each group

	MTX group, n = 85	SSZ group, n = 258	p Value	Difference between mean changes, adjusted for baseline*	
				Unadjusted for propensity	Adjusted for propensity
Continuous variables					
SJC: median change (IQR)	-3 (-9, 0)	-4 (-11, 0)	0.3	Mean (95% CI) -1.12 (-2.78 to 0.54)	Mean (95% CI) -0.59 (-2.64 to 1.45)
TJC: median change (IQR)	-2 (-11, 0)	-4 (-10, 1)	0.9	-2.66 (-5.11 to -0.20)	-1.33 (-4.36 to 1.97)
HAQ: median change (IQR)	-0.3 (-0.7, 0.1)	-0.3 (-0.6, 0.1)	0.9	0.00 (-0.18 to 0.17)	-0.03 (-0.25 to 0.18)
Categorical variables					
Remission: number (%)	5 (6)	14 (5)	0.9	OR (95% CI) 1.1 (0.4 to 3.1)	OR (95% CI) 0.8 (0.2 to 2.8)
Still on same treatment at 2 years: number (%)	48 (56)	129 (50)	0.3	1.3 (0.8 to 2.1)	1.8 (0.9 to 3.3)
Died: number (%)†	4 (4)	7 (2)	0.4		

HAQ, Health Assessment Questionnaire; IQR, interquartile range; MTX, methotrexate; SJC, swollen joint count; SSZ, sulfasalazine; TJC, tender joint count.

*This was done using linear regression, with the change in outcome as the dependent variable, the baseline value of the outcome as a continuous predictor and the group membership (SSZ or MTX) as an indicator variable. The value in the table is the coefficient of the indicator variable.

†Proportion of deaths refers to initial cohort.

the categorisation as rheumatoid arthritis, using the current criteria, is unstable in early arthritis¹⁷ and patients do not necessarily satisfy criteria for rheumatoid arthritis at baseline and in real life; rheumatologists do not necessarily wait for criteria to be satisfied before starting DMARD treatment.

There are, however, some important disadvantages. Firstly, an observational study, even after adjusting for treatment propensity, cannot allow for unmeasured confounders that might have influenced treatment allocation.¹⁸ The distribution of baseline characteristics between the two groups in each of the propensity stratum provided some reassurance that no major confounders were missed. Secondly, as patients were reviewed only annually, there are no robust data on some key variables that may have influenced the outcome that would be normally collected in a well-conducted clinical trial, such as the use of cotherapies such as non-steroidal anti-inflammatory drugs or steroid injections. It is interesting to note that most of the patients stayed with their original drug. We also had no robust data on side effects and their timing. The mortality was slightly higher in the MTX group (10%) than in the SSZ group (7%), but this reflects the slightly higher age at onset (table 1).

MTX has been described as “the anchor drug” for the treatment of early rheumatoid arthritis¹⁹ and it is probably the most widely used preferred DMARD, particularly in North America. It has been reported from more selected cohorts that patients stay on MTX longer than on other DMARDs,²⁰ as

was the case in this study. SSZ is, however, widely used in Europe and the rest of the world. Surprisingly, data on only a few clinical trials are available on the relative efficacy of these two agents. There have been two published trials,^{6,9} whose main focus was different—that is, to compare the results with these two agents when given singly and when given in combination. From the data in these trials, there was no difference between the performance of the drugs, although the studies were only of 1-year duration and, unlike this study, were not able to examine radiological outcomes at 5 years. These trials were also small in size. The first trial recruited just <70 patients in each treated group but had 80% power to detect an effect size of 0.5.⁶ By contrast, the second trial recruited only 35 patients in each group and was almost certainly underpowered to detect a meaningful difference, although the power was not stated in the paper.⁹ Our study was obviously much larger, but real differences may have been missed between the drugs because of small sample size. It is not possible to calculate statistical power in the conventional sense by using the propensity approach, but the 95% CI for the propensity-adjusted differences gives an indication of the probable range of the true difference between the drugs.

We conclude that both SSZ and MTX are effective agents, but there is good evidence, despite the shortfalls of an observational approach, that MTX has a much stronger potential for suppressing the development of erosions, which would support its being the preferred DMARD.

Table 3 Five-year outcome for each group

	MTX group, n = 86	SSZ group, n = 273	p Value	Difference between mean changes, adjusted for baseline	
				Unadjusted for propensity	Adjusted for propensity
Continuous variables					
SJC: median change (IQR)	-4 (-9, 1)	-6 (-13, -1)	0.02	Mean (95% CI) 0.73 (-0.46 to 1.93)	Mean (95% CI) -0.75 (-2.20 to 0.69)
TJC: median change (IQR)	-3 (-11, 2)	-6 (-16, -1)	0.01	2.50 (0.24 to 4.75)	-0.00 (-2.75 to 2.75)
CRP*: median change (IQR)	-6.5 (-28, 5)	-5 (-21, 3)	0.7	6.08 (-1.50 to 13.66)	4.02 (-4.50 to 12.54)
DAS28*: median change (IQR)	-1.4 (-2.7, -0.1)	-1.4 (-2.3, -0.4)	0.9	-0.07 (-0.61 to 0.46)	-0.06 (-0.69 to 0.57)
HAQ: median change (IQR)	-0.1 (-0.6, 0.4)	0 (-0.6, 0.5)	0.4	-0.06 (-0.26 to 0.13)	-0.10 (-0.34 to 0.15)
Larsen Score: median (IQR)	8 (0, 28)	14 (2, 36)	0.3	-1.5 (-44% to 28%)	-31 (-59% to 15%)
Categorical variables					
Erosive†: number (%)	24 (59)	117 (63)	0.6	OR (95% CI) 0.8 (0.4 to 1.7)	OR (95% CI) 0.3 (0.1 to 0.8)
Remission: number (%)	12 (14)	40 (15)	0.4	1.0 (0.5 to 2.0)	1.0 (0.4 to 2.4)
Still on same treatment at 5 years: number (%)	29 (34)	62 (22)	0.04	1.7 (1.0 to 2.9)	2.2 (1.1 to 4.5)
Died: n (%)	11 (10)	24 (7)	0.3		

CRP, C reactive protein; DAS, Disease Activity Score; HAQ, Health Assessment Questionnaire; IQR, interquartile range; MTX, methotrexate; SJC, swollen joint count; SSZ, sulfasalazine; TJC, tender joint count.

*38 patients treated with MTX and 171 patients treated with SSZ had CRP at baseline and 5 years.

†41 patients treated with MTX and 186 patients treated with SSZ had x rays.

ACKNOWLEDGEMENTS

The NOAR is funded by the UK Arthritis Research Campaign. The substantial support of Professor David Scott and clinical colleagues at the Norfolk and Norwich Hospital and the local primary care physicians in recruiting patients and providing access to clinical data is gratefully acknowledged.

Authors' affiliations

S L Hider, A Silman, S Manning, D Symmons, M Lunt, Arc Epidemiology Unit, Manchester University, Manchester, UK

D Bunn, Norfolk Arthritis Register, Norfolk and Norwich University Hospital, Norwich, UK

Funding: This study was funded by the Arthritis Research Campaign UK.

Competing interests: None declared.

REFERENCES

- Fries JF, Williams CA, Morfeld D, Singh G, Sibley J. Reduction in long-term disability in patients with rheumatoid arthritis by disease-modifying antirheumatic drug-based treatment strategies. *Arthritis Rheum* 1996;**39**:616–22.
- Aletaha D, Smolen JS. The rheumatoid arthritis patient in the clinic: comparing more than 1300 consecutive DMARD courses. *Rheumatology* 2002;**41**:1367–74.
- Jobanputra P, Wilson J, Douglas K, Burls A. A survey of British rheumatologists' DMARD preferences for rheumatoid arthritis. *Rheumatology* 2004;**43**:206–10.
- Felson DT, Anderson JJ, Meenan RF. The comparative efficacy and toxicity of second-line drugs in rheumatoid arthritis. Results of two metaanalyses. *Arthritis Rheum* 1990;**33**:1449–61.
- 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. European Leflunomide Study Group. *Lancet* 1999;**353**:259–66.
- Dougados M, Combe B, Cantagrel A, Goupille P, Olive P, Schattenkirchner M, *et al*. Combination therapy in early rheumatoid arthritis: a randomised, controlled, double blind 52 week clinical trial of sulphasalazine and methotrexate compared with the single components. *Ann Rheum Dis* 1999;**58**:220–5.
- Strand V, Cohen S, Schiff M, Weaver A, Fleischmann R, Cannon G, *et al*. Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate. Leflunomide Rheumatoid Arthritis Investigators Group. *Arch Intern Med* 1999;**159**:2542–50.
- Bathon JM, Martin RW, Fleischmann RM, Tesser JR, Schiff MH, Keystone EC, *et al*. A comparison of etanercept and methotrexate in patients with early rheumatoid arthritis. *N Engl J Med* 2000;**343**:1586–93.
- Haagsma CJ, van Riel PL, de Jong AJ, van de Putte LB. Combination of sulphasalazine and methotrexate versus the single components in early rheumatoid arthritis: a randomized, controlled, double-blind, 52 week clinical trial. *Br J Rheumat* 1997;**36**:1082–8.
- Maillefert JF, Combe B, Goupille P, Cantagrel A, Dougados M. Long term structural effects of combination therapy in patients with early rheumatoid arthritis: five year follow up of a prospective double blind controlled study. *Ann Rheum Dis* 2003;**62**:764–6.
- Wiles NJ, Lunt M, Barrett EM, Bukhari M, Silman AJ, Symmons DP, *et al*. Reduced disability at five years with early treatment of inflammatory polyarthritis: results from a large observational cohort, using propensity models to adjust for disease severity. *Arthritis Rheum* 2001;**44**:1033–42.
- Symmons DP, Barrett EM, Bankhead CR, Scott DG, Silman AJ. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. *Br J Rheumatol* 1994;**33**:735–9.
- Flendrie M PFJ. 2004. Calculating the DAS28 using CRP. <http://www.das-score.nl/www.das-score.nl/dascalculators.xls> (accessed 15 Dec 2004).
- 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.
- Bukhari M, Harrison B, Lunt M, Scott DGI, Symmons DPM, Silman AJ. Time to first occurrence of erosions in inflammatory polyarthritis—results from a prospective community-based study. *Arthritis Rheum* 2001;**44**:1248–53.
- Hosmer DW, Lemeshow S. A goodness of fit test for the multiple logistic regression model. *Commun Stat* 1980;**A10**:1043–69.
- Wiles N, Symmons DPM, Harrison B, Barrett E, Barrett JH, Scott DGI, *et al*. Estimating the incidence of rheumatoid arthritis—trying to hit a moving target? *Arthritis Rheum* 1999;**42**:1339–46.
- Landewe RBM. The benefits of early treatment in rheumatoid arthritis: confounding by indication, and the issue of timing. *Arthritis Rheum* 2003;**48**:1–5.
- Pincus T, Yazici Y, Sokka I, Aletaha D, Smolen JS. Methotrexate as the “anchor drug” for the treatment of early rheumatoid arthritis. *Clin Exp Rheumatol* 2003;**21**:S179–85.
- Maetzel A, Wong A, Strand V, Tugwell P, Wells G, Bombardier C. Meta-analysis of treatment termination rates among rheumatoid arthritis patients receiving disease-modifying anti-rheumatic drugs. *Rheumatology* 2000;**39**:975–81.