

EXTENDED REPORT

Survival and effectiveness of leflunomide compared with methotrexate and sulfasalazine in rheumatoid arthritis: a matched observational study

D Aletaha, T Stamm, T Kapral, G Eberl, J Grisar, K P Machold, J S Smolen

Ann Rheum Dis 2003;**62**:944–951

See end of article for authors' affiliations

Correspondence to:
Dr D Aletaha, Division of Rheumatology, Department of Internal Medicine III, University of Vienna, Vienna General Hospital, Waehringer Guertel 18–20, A-1090 Vienna, Austria;
daniel.aletaha@akh-wien.ac.at

Accepted 5 March 2003

Objective: To determine the survival and clinical effectiveness of leflunomide (LEF) compared with methotrexate (MTX) and sulfasalazine (SSZ) for RA in an observational study.

Methods: An observational database of 1088 patients and 5141 patient years of DMARD treatment (2680 courses) from two academic hospitals was filtered for treatment with LEF, MTX, and SSZ. LEF treatment groups were matched for patients' age, baseline ESR, number of previous DMARDs, and hospital cohort with MTX and SSZ treatment groups. For these treatments, Kaplan-Meier analyses of time until the drug was discontinued (drug "survival"), and the effectiveness and safety of continuation of treatment, were performed. The change in disease activity markers (CRP, ESR) was compared between the groups.

Results: The median dose during the study increased from 10 to 15 mg MTX/week and from 1.5 to 2.0 g SSZ/day. Matched survival analysis showed better retention rates for MTX (mean (SEM) survival 28 (1) months) than for LEF (20 (1) months; $p=0.001$), whereas retention rates of SSZ (23 (1) months) were similar to those of LEF ($p=NS$). Treatments were stopped earlier because of adverse events (AEs, 3 months) than because of ineffectiveness (IE, 10 months; $p<0.001$). LEF and MTX were less likely to be stopped because of AEs than SSZ. LEF courses were stopped earlier for AEs ($p<0.001$) than MTX.

Conclusions: Current dosing strategies should be re-evaluated, and coping strategies for common AEs should be investigated. This will be necessary to achieve better drug retention of LEF. At present, MTX continues to be the most effective drug in clinical practice.

Traditional disease modifying antirheumatic drugs (DMARDs), such as methotrexate (MTX) or sulfasalazine (SSZ), form the basis of most treatments for rheumatoid arthritis (RA). In many patients they retard, and sometimes even halt, progression of the disease.^{1–3} However, the long term outcomes with DMARDs are still not satisfactory.^{6–7} Because high standards of effectiveness apply in clinical practice and tolerability of DMARDs is a major prerequisite,^{8–10} the average length of treatment is short, and a series of different DMARDs have to be used during the course of the disease.^{11–13} At the end of the past decade, new DMARDs were licensed and provided new opportunities for the treatment of refractory disease. Like MTX over the previous 10–15 years, with growing clinical experience, these DMARDs are more and more used earlier in the course of disease to effectively prevent disease progression.

One of these new drugs is leflunomide (LEF), a de novo pyrimidine synthesis inhibitor. Its targets of action are lymphocyte activation, cell migration, and activation of transcription factor NF- κ B, which are supposed to have key roles in the pathogenesis of RA.^{14–15} Clinical trials of LEF have provided clear evidence that the signs and symptoms of disease, and of radiographic progression, are reduced^{3–4–16–17} and evidence that the decline of function¹⁸ in patients with RA is prevented. For the management of patients with RA, however, it is also important to know about the long term outcomes with different DMARDs,¹⁹ especially the potential limitations when trial data are transposed to clinical practice. This is an area in which observational studies are invaluable for implementing the results of randomised trials in clinical practice.^{20–21} This study is based on such an observational dataset; its data on DMARD treatments for RA were collected and recorded prospectively at the patient visits. We aim to determine the retention of LEF, MTX, and SSZ, and the ability of these drugs to reduce disease activity in patients with RA.

PATIENTS AND METHODS

Patients and treatments

The basis of this study is an observational dataset of treatments with DMARDs for RA at two rheumatology outpatient clinics in Vienna, the General Hospital and the Lainz Hospital. The database was started in 1999, recording data on all DMARD treatments, such as the time of starting and ending treatment, reason for discontinuation, as well as C reactive protein (CRP), erythrocyte sedimentation rate (ESR), and joint counts, prospectively; in addition, data from a retrospective database with information on treatments back to 1980 were available.^{13–22} At the start of the study the total database comprised 1088 patients and 5141 patient years of DMARD treatment (2680 courses). Fulfilment of the American College of Rheumatology classification criteria for rheumatoid arthritis²³ was a prerequisite for patient inclusion in the database. Figure 1 gives details of patient characteristics and treatments. For this study, only courses of treatment with LEF (n=168), MTX (n=834), and SSZ (n=447) were analysed.

Study end points

The primary outcome was the time until drug discontinuation ("survival of drug", fig 1), which reflects both, the clinical effectiveness and the absence of adverse events (AEs).^{13–24–25} Secondary end points were changes in CRP and ESR measured at the start and end of treatment or at the last visit. Changes

Abbreviations: AEs, adverse events; CRP, C reactive protein; ESR, erythrocyte sedimentation rate; DMARDs, disease modifying antirheumatic drugs; LEF, leflunomide; MTX, methotrexate; RA, rheumatoid arthritis; SSZ, sulfasalazine

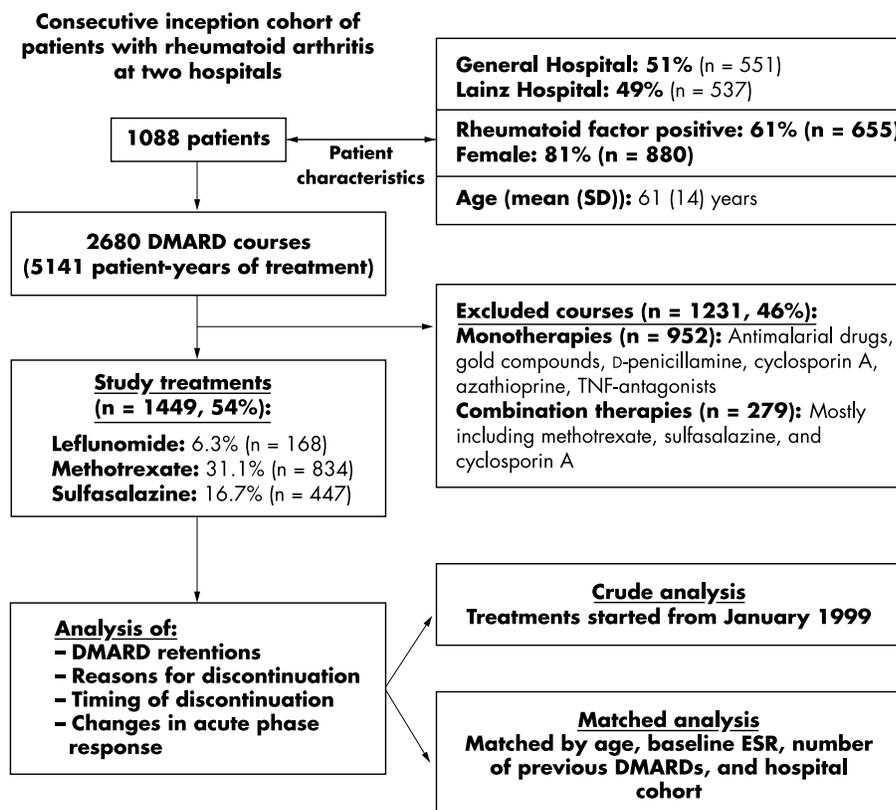


Figure 1 Population studied, characteristics, and flow chart of analysis. The study group comprised a consecutive inception cohort of 1088 patients with RA, who had undergone 2680 DMARD courses.

in tender and swollen joint counts were not used as end points because many of the comparator treatments (MTX, SSZ) were used before 1999 when detailed joint counts were not documented prospectively. However, CRP and ESR are established markers of inflammation and disease activity,²⁶⁻²⁸ which are precise and do not depend on observer judgment. These outcomes were compared for LEF v MTX and LEF v SSZ. Two different analyses were performed to assess these comparisons (fig 1): (a) looking generally at all treatments after 1999 (crude analysis, no matching); (b) matching LEF courses with MTX and SSZ courses to adjust for bias by indication (matched analysis)—that is, the fact that the choice of DMARD regimens is affected by patient and disease characteristics before treatment. In the latter, SSZ and MTX courses before 1999 were included to optimise the availability of matches. The matching algorithm is described below.

Predictors of drug survival

The “survival” of a DMARD as the primary end point of this study was determined using the Kaplan-Meier estimator, which can deal with so-called “censored” cases. These contribute to the analysis by the length of their period of observation, but treatment discontinuation (as the end point for the analysis) has not yet occurred at the time of data evaluation. The equivalence of survival distributions of the different DMARDs was statistically assessed with the Breslow test, which weights the comparators by the proportion of subjects at risk at different times.²⁹

To match treatments with a similar likelihood of being discontinued, we tested potential predictors of drug survival by a stepwise Cox regression model. Age and the total number of previous DMARD courses in a single patient were strong predictors of drug survival in this model, with increased likelihood of discontinuation for lower age ($p=0.002$, Wald statistics) and for a higher number of previous DMARDs ($p=0.0001$). Baseline levels of acute phase reactants were only weakly associated with increased likelihood of drug discontinuation, with ESR contributing more ($p=0.028$) than CRP

($p=NS$) in the model. The hospital cohort (General Hospital or Lainz Hospital) was also a weak predictor, but it was also incorporated in the matching algorithm (see later), to dismiss the potential influence of different practices at different hospitals. Further characteristics tested and not associated with DMARD retention were sex, rheumatoid factor status, and time to DMARD initiation (from first onset of symptoms). We also tested for a possible interaction of the three variables: age, time before a DMARD was used, and total number of DMARDs on the outcome variable. The regression model showed no effect modification as we included the three interaction terms.

Matching algorithm

We matched LEF treatments with MTX and SSZ treatments in two steps: first, we grouped DMARD courses by hospital cohort (General or Lainz Hospital), and then subdivided these further into three groups according to the number of previous DMARDs (no DMARD, 1 DMARD, >1 DMARD), resulting in six groups; in a second step, we determined the best matching MTX or SSZ courses for each LEF course within the same group by the minimum Euclidian distances for the variables age and ESR.

Adjustment for time at risk

We found no temporal trend in DMARD retention when we included the year of application in a Cox regression model. This was also the case in a backup analysis, in which Kaplan-Meier plots for DMARD retention (of MTX and SSZ) according to the year of application (using two year categories) showed similar slopes ($p=NS$, Breslow test; data not shown).

In the Kaplan-Meier model treatments like MTX and SSZ may be favoured because of their potential to achieve treatment for a longer period of time during the two decades of clinical use, compared with LEF, which has been used only since 1999. Therefore, LEF courses have a shorter duration simply as a consequence of their more recent entry into the therapeutic armament. We adjusted for this difference by

Table 1 Patient and disease characteristics at treatment initiation of leflunomide (LEF), methotrexate (MTX), and sulfasalazine (SSZ). (A) Unmatched data of all available treatments. (B) After matching LEF treatments with MTX and SSZ treatments, respectively

| A | LEF (n=168) | MTX (n=834) | SSZ (n=447) | |
|--|----------------|----------------|----------------|--|
| Cohort, % General Hospital | 62.5 | 54.6 | 56.4 | |
| Sex, % female | 82.5 | 81.1 | 80.9 | |
| Rheumatoid factor status, % positive | 69.6 | 69.4 | 59.3* | |
| Age (years), mean (SD) | 58.4 (12.9) | 61.1 (13.6) | 59.8 (14.8) | |
| Previous DMARDs, median number (quartiles) | 3 (2; 4)† | 1 (0; 2) | 1 (0; 2) | |
| ESR (mm/1st h), mean (SD) | 38 (27) | 42 (27) | 41 (26) | |
| CRP (mg/l), mean (SD) | 26.2 (29.3) | 31.0 (31.2) | 25.3 (26.8)‡ | |

* χ^2 Pearson: $p=0.002$; †Mann-Whitney U test: $p<0.001$ compared with MTX and SSZ; ‡ t test: $p=0.002$ compared with MTX.

| B | MTX matches* | | SSZ matches* | |
|--|----------------|----------------|----------------|----------------|
| | LEF (n=168) | MTX (n=168) | LEF (n=159) | SSZ (n=159) |
| Cohort, % General Hospital | 62.5 | 59.5 | 64.8 | 56.0 |
| Sex, % female | 82.5 | 78.6 | 81.5 | 81.8 |
| Rheumatoid factor status, % positive | 69.6 | 71.9 | 69.1 | 61.2 |
| Age (years), mean (SD) | 58.4 (12.9) | 58.4 (12.3) | 58.6 (12.9) | 60.1 (13.0) |
| Previous DMARDs, median number (quartiles) | 3 (2; 4) | 2 (2; 4) | 3 (2; 4) | 1 (0; 3)† |
| ESR (mm/1st h), mean (SD) | 38 (27) | 38 (27) | 39 (27) | 38 (26) |
| CRP (mg/l), mean (SD) | 26.2 (29.3) | 27.3 (27.4) | 25.8 (27.7) | 22.7 (22.3) |

*All matched pairs were derived from the same hospital cohort; †Mann-Whitney U test: $p<0.001$ compared with LEF.

introducing four 42 month periods of application (July 1988–December 1991; January 1992–June 1995; July 1995–December 1998; January 1999–June 2002), assuming that the use of MTX and SSZ during all periods was similar and comparable with that of LEF in the most recent group. If treatments were still continuing at the arbitrary end points of these periods, they were censored for the analysis (as described above); similarly, current treatments were censored in the last period.

Statistical analyses

We analysed the timing of inefficiency and AEs, and used the Kolmogorov-Smirnov test to compare the temporal distribution between the three DMARDs.³⁰ The Wald test and Breslow test statistics were used as mentioned above. For the comparison of DMARD dosage over the four periods we used the Kruskal-Wallis test, a generalisation of the Mann-Whitney U test for non-parametric data and comparison of more than two groups. Finally, we assessed DMARD effectiveness between the matched treatments by comparing their ability to reduce the markers of disease activity (t test). For matching and outcome analysis we used the statistical package for the social sciences, version 11 (SPSS, Chicago, IL, USA).

RESULTS

Characteristics in the different subgroups

Table 1A shows that sex, rheumatoid factor status, and age were similar in the different therapeutic subgroups except that rheumatoid factor was positive in fewer patients in the SSZ group than in the MTX and LEF groups. CRP and ESR were highest at the start of MTX treatments. After matching the therapeutic segments no significant difference between the treatment groups was seen (table 1B). Only the number of previous DMARDs was still significantly higher in LEF compared with SSZ treatments, but the proportions in the matching groups (no previous DMARD/1 DMARD/>1 previous DMARD) were identical. It has been shown previously that the association between the number of previous courses and treatment duration is strongest during the first few

courses.^{31–32} MTX courses that were matched with LEF courses showed lower mean values of CRP and ESR at the start of treatment than all courses of MTX. This indicates that LEF was given to patients with more refractory rather than more active disease. For nine LEF treatments no SSZ match could be found. The median dose of DMARD increased significantly over the years of application (Kruskal-Wallis test: $p<0.001$), from a median of 10 mg/week to 15 mg/week for MTX, and 1.5 g/day to 2.0 g/day for SSZ (table 2).

Comparative survival of LEF

The maximum possible treatment duration of LEF in this study is 42 months because it was introduced into clinical practice in these hospitals in January 1999. To allow comparison of treatment durations with MTX or SSZ, for which the period of observation is much longer, 42 month intervals were introduced and continuing treatments were censored accordingly (see “Patients and methods”) before performing survival estimates. The Kaplan-Meier plots (fig 2) are presented according to recent recommendations.³³ A vertical line indicates a disproportionately small number (<10%) of original patients at risk.

Crude analysis

First, we estimated the survival of all available treatments with LEF, MTX, and SSZ beginning January 1999 or later. This

Table 2 Median doses (and quartiles) of methotrexate (MTX) and sulfasalazine (SSZ) over the years (42 month periods)*

| | MTX (mg/week) | SSZ (g/day) |
|-----------------------|----------------|----------------|
| (1) Jul 1988–Dec 1991 | 10 (7.5; 10) | 1.5 (1.0; 3.0) |
| (2) Jan 1992–Jun 1995 | 10 (7.5; 12.5) | 2.0 (1.5; 2.0) |
| (3) Jul 1995–Dec 1998 | 12.5 (10; 15) | 2.0 (1.5; 3.0) |
| (4) Jan 1999–Jun 2002 | 15 (10; 20) | 2.0 (2.0; 3.0) |

*Kruskal-Wallis-test: $p<0.001$ for dose changes over time periods for both MTX and SSZ. Median doses of the total cohort: MTX 12.5 mg/week (10; 15); SSZ 2.0 g/day (1.5; 3.0).

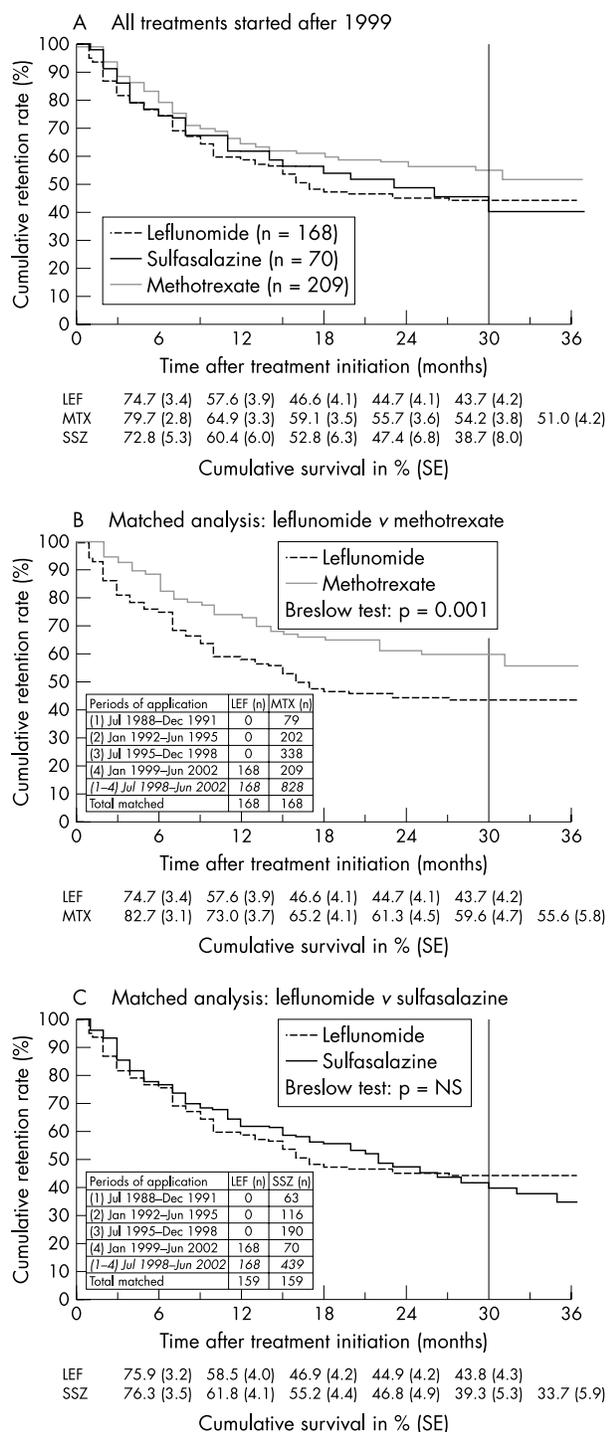


Figure 2 Cumulative drug retentions (%) of LEF, MTX, and SSZ. The period to the right of the reference line (at 30 months) indicates an unduly small number of patients left at risk (<10%). (A) All treatments started after 1999, mean survivals (SE) of the drugs (months)—MTX (n=209): 26 (1); SSZ (n=70): 23 (2); LEF (n=168): 20 (1). Breslow test: p=0.03 for MTX v LEF, otherwise: p=NS. (B) Matched analysis: LEF (n=168) v MTX (n=168)—MTX: 28 (1) months; LEF: 20 (1) months. Breslow test: p=0.001. (C) Matched analysis: LEF (n=159) v SSZ (n=159)—LEF: 20 (1) months; SSZ: 23 (1) months. Breslow test: p=NS.

was done regardless of any potential baseline differences between the treatment groups. Figure 2A shows that the mean (SE) drug retention for MTX (26 (1) months) was significantly better than for LEF (20 (1) months) (Breslow test:

p=0.03), whereas the curves were similar for SSZ (23 (2) months) and LEF treatments (p=NS).

Matched analysis

Next we used the algorithm described above to match LEF courses with MTX and SSZ courses which had similar a priori values of predictors of treatment duration. The matched analysis (fig 2B) emphasises further (p=0.001) the better retention of MTX (28 (1) months) compared with that of LEF (20 (1) months) that was found in the crude analysis. As mentioned before, it seems that, in general, LEF was not given to patients with higher disease activity, but rather to those with more refractory disease (table 1). Again, the curves of LEF (20 (1)) and SSZ (23 (1)) are similar (fig 2C, p=NS).

Reasons for, and timing of, treatment discontinuation

Reasons
In the 42 month treatment groups, treatments were ended mainly because of insufficient effectiveness (IE) or AEs. Other reasons for discontinuation included non-compliance (treatment ended by patient), surgical procedures, (planned) pregnancy, or concurrent comorbidity (impeding further drug ingestion). Examination of the crude data for the first 42 months of treatment shows that LEF treatments were stopped more often because of inefficiency, and less often because of AEs, than MTX and SSZ (p=0.02, χ^2 test; table 3A). After matching for baseline characteristics these differences decreased (p=NS, χ^2 test; table 3B).

Timing

Inefficiency (IE) limits continuation of DMARD treatments after a median of 10 months of treatment (quartiles: 6; 18 months, fig 3A), adverse events (AE) do so after a median of 3 months (1; 10 months) (fig 3B). Because the follow up of the treatment courses is arbitrarily cut after 42 months, these descriptives (median, quartiles) are likely to be underestimated. In consequence, for statistical comparison of these findings we used the Kolmogorov-Smirnov test. This test examines the similarity of the event distribution within the observed time frame, and was highly significant for IE versus AE regardless of the DMARD employed (p<0.001). If the timing of discontinuations is compared for the three different DMARDs, the curves of inefficiency (fig 3A) indicate similar timing (median and quartiles) for discontinuation of MTX (11 and 6; 19 months) and SSZ (10 and 5; 19 months) and earlier discontinuation of LEF (7 and 5; 15 months); the Kolmogorov-Smirnov test was not significant for this finding. In contrast, treatment was ended significantly earlier for AEs when LEF (3 and 1; 7 months) or SSZ (3 and 1; 7 months) was used compared with MTX (6 and 2; 14 months) (p<0.001; fig 3B).

Survival of safety and effectiveness

Figures 4 and 5 illustrate an additional approach to the problem of sequential event determination. Again, we used the Kaplan-Meier method to determine the “survival” of effectiveness (fig 4) and safety (fig 5) of the different drugs. For this purpose, only discontinuations for inefficiency (fig 4) are counted as an event, while all other discontinuations contribute to the “no event observed” group and are censored at the respective times. Conversely, in fig 5, only discontinuations due to AEs are counted. For both analyses, we used two models: firstly, using only treatments beginning January 1999 or later (figs 4A and 5A); and secondly, using the matched treatments as described above comparing LEF with MTX (figs 4B and 5B) and with SSZ (figs 4C and 5C).

For both, effectiveness and safety, it can be seen that the MTX and SSZ curves approach that of LEF after matching (figs 4 and 5). This simply reflects the successful adjustment for potential bias. However, if effectiveness were the single determinant of treatment duration, which is a hypothetical

Table 3 Reasons for stopping excluded methotrexate (MTX), sulfasalazine (SSZ), and leflunomide (LEF) treatments. * (100% = all excluded treatments with the respective regimen) (A) Unmatched data using all available treatments; χ^2 test, $p=0.02$. (B) Matched treatments; χ^2 test, $p=NS$. Results are shown as percentages

| A | MTX (n=274) | SSZ (n=195) | LEF (n=84) | Total (n=553) |
|-------------------|--------------|-------------|-------------|---------------|
| Inefficiency | 39.1 (n=107) | 43.6 (n=85) | 54.8 (n=46) | 43.0 (n=238) |
| Adverse events | 37.6 (n=103) | 41.5 (n=81) | 27.4 (n=23) | 37.4 (n=207) |
| Lost to follow up | 10.9 (n=30) | 7.2 (n=14) | 13.1 (n=11) | 9.9 (n=55) |
| Other reasons* | 12.4 (n=34) | 7.7 (n=15) | 4.8 (n=4) | 9.6 (n=53) |

| B | MTX matches | | SSZ matches | |
|-------------------|-------------|-------------|-------------|-------------|
| | MTX | LEF | SSZ | LEF |
| Inefficiency | 50.0 (n=21) | 54.8 (n=46) | 44.8 (n=30) | 55.7 (n=44) |
| Adverse events | 28.6 (n=12) | 27.4 (n=23) | 31.3 (n=21) | 26.6 (n=21) |
| Lost to follow up | 9.5 (n=4) | 13.1 (n=11) | 11.9 (n=8) | 12.7 (n=10) |
| Other reasons† | 11.9 (n=5) | 4.8 (n=4) | 11.9 (n=8) | 5.1 (n=4) |

*Contributions may not exactly add up to 100% owing to rounding errors; †includes non-compliance, concurrent comorbidity, surgery, pregnancy.

assumption, then MTX would be maintained significantly longer than LEF (mean survival (SE) in months: 32 (1) and 26 (1), respectively; $p=0.04$, Breslow test) (fig 4B). The difference between the “survival” of matched SSZ and LEF treatments (30 (1) and 26 (1), respectively) was not significant. If treatments were withdrawn only for AEs, then, in accordance

with the former analyses, there would be no major drop in the curves after 6–12 months. When matched courses were compared for survival of safety (figs 5B and C), there was no significant difference in the hypothetical survival times between MTX (34 (1) months), SSZ (33 (1) months), and LEF (31 (1) months).

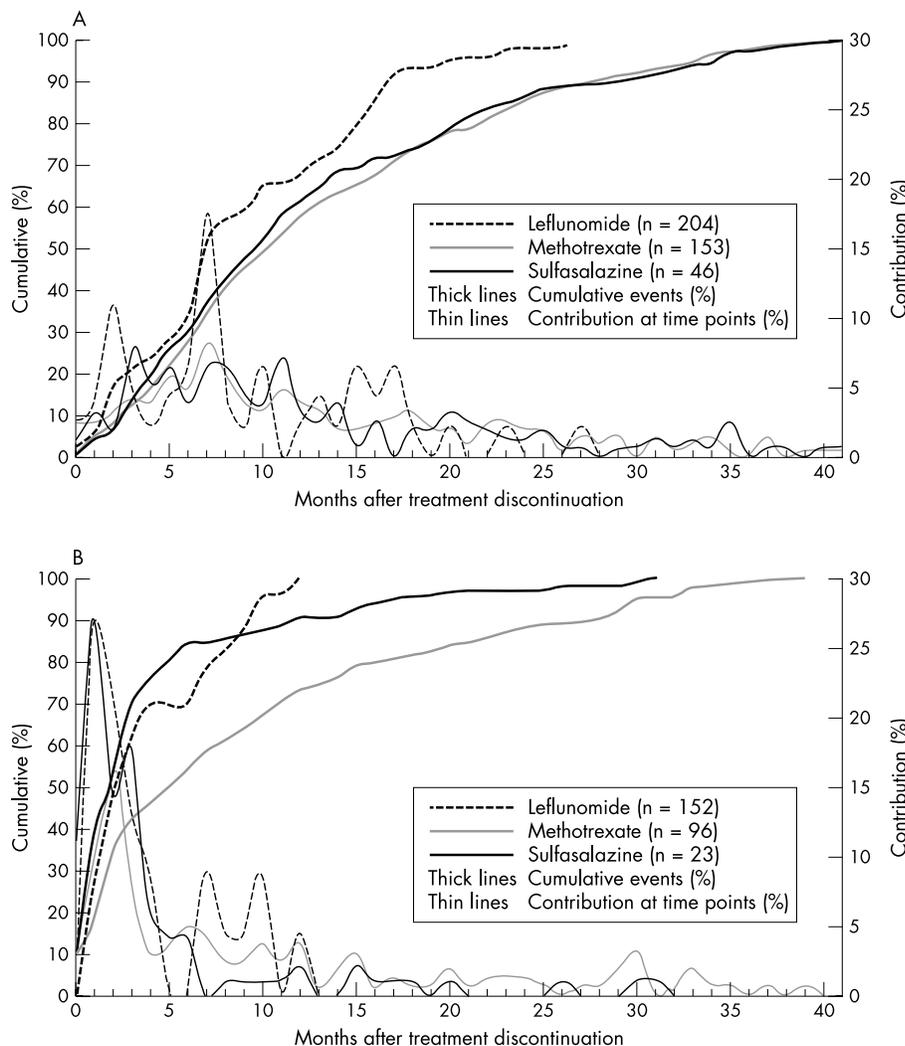


Figure 3 Timing of DMARD discontinuation due to (A) inefficiency (A) and (B) AEs. Only treatments with an event during the first 42 months are displayed (n). The median timing of discontinuations (for all DMARDs) due to inefficiency (A) was 10 months (quartiles: 6; 18 months); and due to adverse events (B) was: median 3 months (1; 10 months) ($p<0.001$, Kolmogorov-Smirnov test for equality of distributions). Between-drug comparison for timing of inefficiency (A): $p=NS$; and for timing of adverse events (B): $p<0.001$ (earlier timing of events while receiving LEF and SSZ than while receiving MTX treatment).

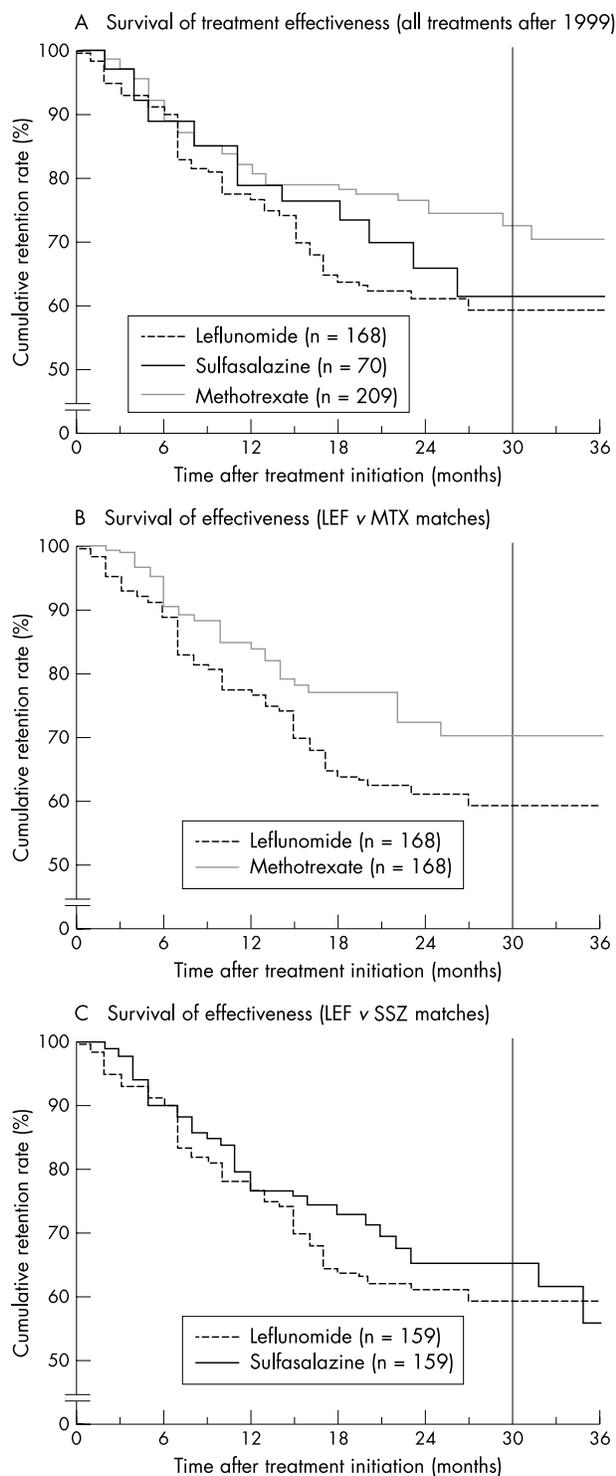


Figure 4 Survival of treatment effectiveness. Cumulative drug retentions (%) of LEF, MTX, and SSZ, when only discontinuation due to inefficiency was analysed, assuming permanent safety otherwise (that is, censoring at time of occurrence). The period to the right of the reference line (30 months) indicates an unduly small number of patients left at risk (<10%). Note the discontinuous y axis. (A) All treatments after 1999, mean survivals (SE) of the drugs (months)—MTX (n=209): 33 (1); SSZ (n=70): 30 (2); and LEF (n=168): 26 (1). Breslow test: $p=0.04$ for MTX v LEF, otherwise: $p=NS$. (B) Matched analysis of LEF v MTX (n=168)—LEF: 26 (1) v MTX: 32 (1); Breslow test: $p=0.04$. (C) Matched analysis of LEF v SSZ (n=159)—LEF: 26 (1) v SSZ: 30 (1); Breslow test: $p=NS$.

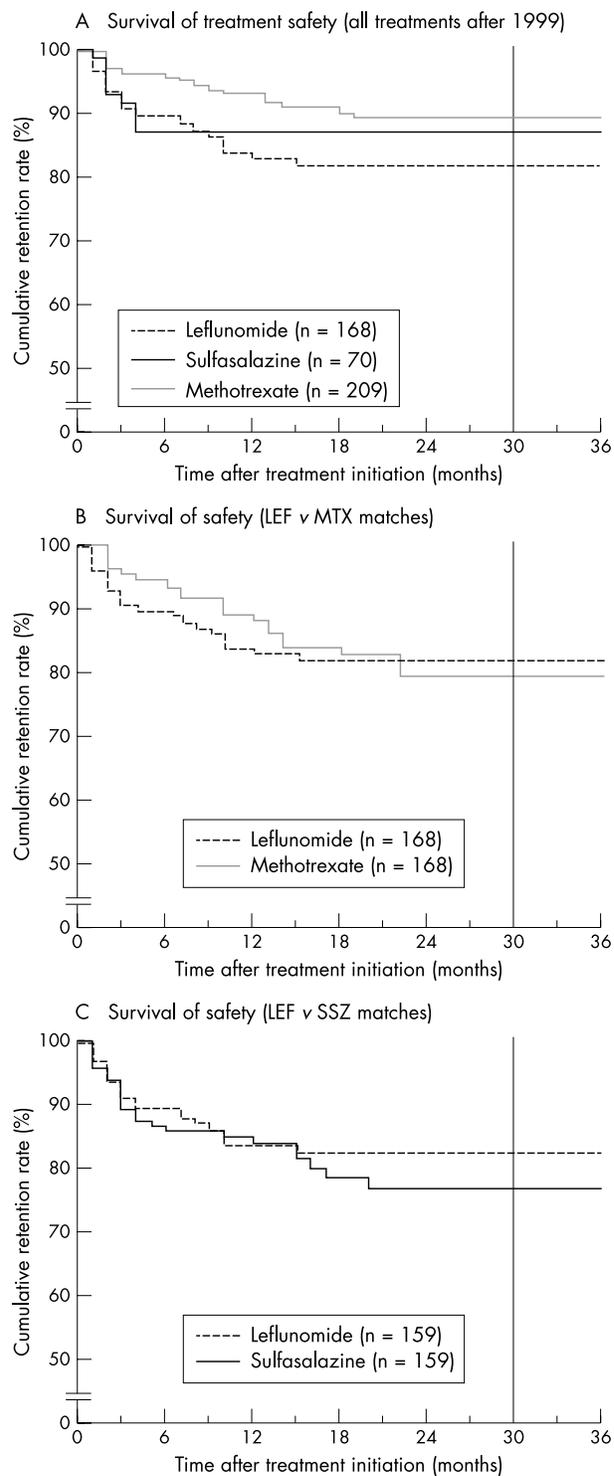


Figure 5 survival of treatment safety. Cumulative drug retentions (%) of LEF, MTX, and SSZ, while only discontinuation due to adverse events was analysed, assuming permanent effectiveness otherwise (that is, censoring at time of occurrence). The period to the right of the reference line (30 months) indicates an unduly small number of patients left at risk (<10%). Note the discontinuous y axis. (A) All treatments after 1999, mean survivals (SE) of the drugs (months)—MTX (n=209): 37 (1); SSZ (n=70): 36 (2); and LEF (n=168): 30 (1). Breslow test: $p=0.01$ for MTX v LEF, otherwise: $p=NS$. (B) Matched analysis of LEF v MTX (n=168)—LEF: 30 (1) v MTX: 34 (1); Breslow test: $p=NS$. (C) Matched analysis of LEF v SSZ (n=159)—LEF: 31 (1) v SSZ: 33 (1); Breslow test: $p=NS$.

Table 4 Improvement of C reactive protein (CRP) and erythrocyte sedimentation rate (ESR) during DMARD treatments. Only matched courses entered into the analysis

| | Reduction of CRP (mg/l) | | | Reduction of ESR (mm/1st h) | | |
|---------------------|-------------------------|-------------|------------------|-----------------------------|------------|------------------|
| | No | Mean (SEM) | t Test (p value) | No | Mean (SEM) | t Test (p value) |
| Leflunomide (LEF) | 129 | -6.0 (2.1)* | v MTX: p=0.45 | 146 | -1.4 (1.9) | v MTX: p=0.03 |
| Methotrexate (MTX) | 107 | -8.4 (2.4) | v SSZ: p=0.04 | 146 | -7.4 (2.0) | v SSZ: p=0.14 |
| Sulfasalazine (SSZ) | 98 | 6.0 (7.0) | v LEF: p=0.06 | 137 | -3.3 (2.0) | v LEF: p=0.49 |

*Changes of CRP in the LEF group: p=0.004 (paired t test); ESR: p=NS.

Changes in acute phase response

In a final step we determined the changes of CRP and ESR between the start and end of DMARD treatment (table 4). In a follow up of patients with RA these are valuable measures of systemic inflammation and markers of disease activity.^{26, 27, 34} The greatest mean (SEM) reduction of CRP occurred during MTX treatment (-8.4 (2.4) mg/l) compared with LEF (-6.0 (2.1) mg/l; p=NS, t test), and SSZ (+6.0 (±7.0) mg/l; p=0.04). Treatment with LEF reduced ESR values less (-1.4 (1.9) mm/1st h) than MTX (-7.4 (2.0) mm/1st h; p=0.03) or SSZ (-3.3 (2.0) mm/1st h; p=NS) treatments. LEF significantly reduced baseline CRP values (p=0.004; paired t test), ESR improvement was not significant. It should also be mentioned that LEF tended to decrease ESR less strongly than CRP, and the trend was in the opposite direction for SSZ, which is in accordance with previously published trial data.³

DISCUSSION

The importance of a synthesis of trial and observational data for new treatments has been discussed previously.^{20, 21} Trial data may underestimate the risk for AEs in therapeutic subgroups,³⁵ and a priori exclude subjects who will later be treated in clinical practice, such as patients with mild disease activity or comorbidity. Thus, in contrast with observational data they do not reflect the performance of treatments in daily life. The primary aim of this study was the comparison of DMARD retention in clinical practice. Such analyses have previously been performed for traditional DMARDs,^{12, 13, 25} but LEF had not yet been analysed, because it was licensed in Europe only in 1999.

Drug retention rates reflect the patient's and doctor's satisfaction with a given treatment, but depending on the availability of therapeutic alternatives, the threshold for drug discontinuation may vary with respect to both the wish for more effectiveness and for less toxicity. In recent years, no matter which DMARDs were used, rheumatologists still had biological agents to offer to their patients. This practice related bias is superimposed by a disease related bias affecting drug retention—for example, that patients with a potentially more refractory disease (a history of repeated failure of traditional DMARDs) were more likely to be receive LEF than yet another traditional DMARD. In fact, most patients receiving LEF had previously been treated with MTX, whereas most patients receiving MTX had a DMARD history not including LEF. The potential of LEF as an initial treatment in MTX-naïve patients (and subsequent matching by DMARD sequence) can presently not be investigated observationally. Conversely, MTX retention may be expected to be worse in patients who had already had an unsuccessful course with LEF. In addition, irrespective of the DMARD type this notion is particularly important in the light of previous observations of decreasing effectiveness with increasing number of DMARD courses and increasing disease duration.^{31, 33} We performed a series of statistical adjustments to minimise this type of bias, and to make the different treatments comparable. Employing all these adjustments, we observed that LEF had a similar retention rate to SSZ, but a significantly lower one than MTX. This

observation conflicts with some trial data,¹⁷ and a few explanations for this controversy should be further discussed.

Although the overall incidence of LEF discontinuations due to AEs tended to be lower than those of MTX and SSZ, discontinuations related to toxicity occurred significantly earlier with LEF (and SSZ) than with MTX. A high rate of early discontinuations related to toxicity was apparent during early LEF treatment, comparable with that of SSZ. The loading doses of LEF, which were given at the start of treatment to all patients, may be a potential reason for this finding. Furthermore, for more traditional DMARDs like MTX, symptomatic drugs for frequent AEs (such as nausea, alopecia, stomatitis, liver toxicity, etc.) have been established and in many cases make it possible to maintain an effective treatment.³⁶ Such strategies, consolidated through many years of clinical use of these drugs and learning, have not yet been evaluated, let alone established for LEF.

Apart from early toxicity, discontinuations due to lack of effectiveness also occurred earlier with LEF than with the other DMARDs, even after matching for baseline criteria. Thus, the current dosing scheme of LEF does not lead to increases in dose in poor responders. The better effectiveness of higher doses of MTX and SSZ seen in the recent past and transposed into clinical practice (table 2) suggests that the effectiveness of LEF might also improve with an increased dose.

In conclusion, our data show that in daily life LEF performs as well as SSZ, an observation in accordance with the trial findings.^{3, 37} In contrast, retention of MTX is longer than retention of LEF. This is likely to be attributable to the very recent insights into optimal dosing and optimal coping with AEs. Because LEF has been rigidly used according to manufacturers' and regulatory authority labels, and because toxicity appears to be increased only during the first few months after the start of treatment, the data presented call for a re-evaluation of current loading dose requirements and dose increases in patients with continuing active disease while having good DMARD tolerability. It took over 20 years of MTX use^{1, 38} to arrive at current recommendations—it should not take as long to develop potential optimal dosing strategies for new DMARDs.

Authors' affiliations

D Aletaha, T Stamm, T Kapral, J Grisar, K P Machold, J S Smolen, Division of Rheumatology, Department of Internal Medicine III, University of Vienna, Austria
G Eberl, J S Smolen, Second Department of Medicine, Lainz Hospital, Vienna, Austria

REFERENCES

- Weinblatt ME, Coblyn JS, Fox DA, Fraser PA, Holdsworth DE, Glass DN, et al. Efficacy of low-dose methotrexate in rheumatoid arthritis. *N Engl J Med* 1985;312:818-22.
- Pullar T, Hunter JA, Capell HA. Sulphasalazine in rheumatoid arthritis: a double blind comparison of sulphasalazine with placebo and sodium aurothiomalate. *BMJ* 1983;287:1102-4.
- 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.
- 4 **Sharp JT**, Strand V, Leung H, Hurley F, Loew-Friedrich I. Treatment with leflunomide slows radiographic progression of rheumatoid arthritis: results from three randomized controlled trials of leflunomide in patients with active rheumatoid arthritis. Leflunomide Rheumatoid Arthritis Investigators Group. *Arthritis Rheum* 2000;43:495-505.
 - 5 **Jeurissen ME**, Boerbooms AM, van De Putte LB, Doesburg WH, Lemmens AM. Influence of methotrexate and azathioprine on radiologic progression in rheumatoid arthritis. A randomized, double-blind study. *Ann Intern Med* 1991;114:999-1004.
 - 6 **Pincus T**. Rheumatoid arthritis: disappointing long-term outcomes despite successful short-term clinical trials. *J Clin Epidemiol* 1988;41:1037-41.
 - 7 **Fries JF**. Current treatment paradigms in rheumatoid arthritis. *Rheumatology (Oxford)* 2000;39:30-5.
 - 8 **Furst DE**. Aggressive strategies for treating rheumatoid arthritis: has the case been proven? *Lancet* 2000;356:183-4.
 - 9 **Emery P**. Therapeutic approaches for early rheumatoid arthritis. How early? How aggressive? *Br J Rheumatol* 1995;34:87-90.
 - 10 **Wilske KR**, Healey LA. The need for aggressive therapy of rheumatoid arthritis. *Rheum Dis Clin North Am* 1993;19:153-61.
 - 11 **Wijnands MJ**, van't Hof MA, van Leeuwen MA, van Rijswijk MH, van De Putte LB, van Riel PL. Long-term second-line treatment: a prospective drug survival study. *Br J Rheumatol* 1992;31:253-8.
 - 12 **Hurst S**, Kallan MJ, Wolfe FJ, Fries JF, Albert DA. Methotrexate, hydroxychloroquine, and intramuscular gold in rheumatoid arthritis: relative area under the curve effectiveness and sequence effects. *J Rheumatol* 2002;29:1639-45.
 - 13 **Aletaha D**, Smolen JS. Effectiveness profiles and dose dependent retention of traditional disease modifying antirheumatic drugs for rheumatoid arthritis. An observational study. *J Rheumatol* 2002;29:1631-8.
 - 14 **Breedveld FC**, Dayer JM. Leflunomide: mode of action in the treatment of rheumatoid arthritis. *Ann Rheum Dis* 2000;59:841-9.
 - 15 **Grisar J**, Eselboeck D, Koeller M, Aringer M, Stummvoll G, Pietschmann P, *et al*. Leflunomide reduces transendothelial migration [abstract]. *Arthritis Rheum* 2001;44(suppl):S297.
 - 16 **Emery P**, Breedveld FC, Lemmel EM, Kaltwasser JP, Dawes PT, Gomor 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.
 - 17 **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.
 - 18 **Strand V**, Tugwell P, Bombardier C, Maetzel A, Crawford B, Dorrier C, *et al*. Function and health-related quality of life: results from a randomized controlled trial of leflunomide versus methotrexate or placebo in patients with active rheumatoid arthritis. Leflunomide Rheumatoid Arthritis Investigators Group. *Arthritis Rheum* 1999;42:1870-8.
 - 19 **Pincus T**, Stein CM. Why randomized controlled clinical trials do not depict accurately long-term outcomes in rheumatoid arthritis: some explanations and suggestions for future studies. *Clin Exp Rheumatol* 1997;15(suppl 17):S27-38.
 - 20 **Hawley DJ**, Wolfe F. Are the results of controlled clinical trials and observational studies of second line therapy in rheumatoid arthritis valid and generalizable as measures of rheumatoid arthritis outcome: analysis of 122 studies. *J Rheumatol* 1991;18:1008-14.
 - 21 **Black N**. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996;312:1215-18.
 - 22 **Aletaha D**, Smolen JS. Laboratory testing in rheumatoid arthritis patients taking disease-modifying antirheumatic drugs: clinical evaluation and cost analysis. *Arthritis Rheum* 2002;47:181-8.
 - 23 **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.
 - 24 **Fries JF**. Effectiveness and toxicity considerations in outcome directed therapy in rheumatoid arthritis. *J Rheumatol Suppl* 1996;44:102-6.
 - 25 **Pincus T**, Marcum SB, Callahan LF. Longterm drug therapy for rheumatoid arthritis in seven rheumatology private practices: II. Second line drugs and prednisone. *J Rheumatol* 1992;19:1885-94.
 - 26 **Wolfe F**. Comparative usefulness of C-reactive protein and erythrocyte sedimentation rate in patients with rheumatoid arthritis. *J Rheumatol* 1997;24:1477-85.
 - 27 **van der Heijde DM**, van't Hof MA, van Riel PL, van de Putte LB. Validity of single variables and indices to measure disease activity in rheumatoid arthritis. *J Rheumatol* 1993;20:538-41.
 - 28 **Atkinson JP**. C-reactive protein: a rheumatologist's friend revisited. *Arthritis Rheum* 2001;44:995-6.
 - 29 **Breslow NE**, Edler L, Berger J. A two-sample censored-data rank test for acceleration. *Biometrics* 1984;40:1049-62.
 - 30 **Conover RJ**. *Practical nonparametric statistics*. 3rd ed. New York: Wiley, 1999.
 - 31 **Anderson JJ**, Wells G, Verhoeven AC, Felson DT. Factors predicting response to treatment in rheumatoid arthritis: the importance of disease duration. *Arthritis Rheum* 2000;43:22-9.
 - 32 **Aletaha D**, Smolen J. The rheumatoid arthritis patient in the clinic: comparing more than 1300 consecutive DMARD therapies. *Rheumatology (Oxford)* 2002;41:1367-74.
 - 33 **Pocock SJ**, Clayton TC, Altman DG. Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. *Lancet* 2002;359:1686-9.
 - 34 **Devlin J**, Gough A, Huissoon A, Perkins P, Holder R, Reece R, *et al*. The acute phase and function in early rheumatoid arthritis. C-reactive protein levels correlate with functional outcome. *J Rheumatol* 1997;24:9-13.
 - 35 **Ellis SJ**, Adams RF. The cult of the double-blind placebo-controlled trial. *Br J Clin Pract* 1997;51:36-9.
 - 36 **Morgan SL**, Baggott JE, Vaughn WH, Young PK, Austin JV, Krumdieck CL, *et al*. The effect of folic acid supplementation on the toxicity of low-dose methotrexate in patients with rheumatoid arthritis. *Arthritis Rheum* 1990;33:9-18.
 - 37 **Scott DL**, Smolen JS, Kalden JR, van de Putte LB, Larsen A, Kvien TK, *et al*. Treatment of active rheumatoid arthritis with leflunomide: two year follow up of a double blind, placebo controlled trial versus sulfasalazine. *Ann Rheum Dis* 2001;60:913-23.
 - 38 **Hoffmeister RT**. Methotrexate therapy in rheumatoid arthritis: 15 years experience. *Am J Med* 1983;75:69-73.