Aventis Pharma

ARAVA® (LEFLUNOMIDE)

(HWA 486)

Briefing Document, February 2003

FDA Advisory Committee Meeting

March 5, 2003

TABLE OF CONTENTS

EXECUTIVE SUMMARY AND INTRODUCTION

ARAVA® (leflunomide) is an isoxazole immunomodulatory agent with a unique mechanism of action. It inhibits de novo pyrimidine synthesis by reversibly blocking the enzyme dihydroorotate dehydrogenase (DHO-DH), resulting in antiproliferative effects on activated autoimmune lymphocytes important in the pathogenesis of rheumatoid arthritis (RA).

ARAVA® has been shown in randomized, controlled trials to: (i) reduce the signs and symptoms of active RA; (ii) retard structural joint damage as evidenced by X-ray assessments of erosions and joint space narrowing; and (iii) improve physical function. In every trial, ARAVA® was consistently significantly superior to placebo and was overall comparable to comparator Disease Modifying Anti-Rheumatic Drugs (DMARDs), methotrexate and sulfasalazine.

ARAVA® targets the underlying inflammatory process – rather than just treating symptoms – by inhibiting multiplication of T-cells believed to perpetuate the autoimmune response in RA. It is also effective in treating both early and long-standing disease, as long- and short-term therapy, and regardless of disease severity or previous exposure to other DMARDs. Because of its unique properties and the need for additional DMARD therapies, ARAVA® received priority review by the U.S. Food and Drug Administration (FDA). This does mean that less data was required to receive approval, but that, by regulation, the FDA acted on an expedited track due to the important therapeutic potential offered by ARAVA®.

The New Drug Application (NDA) for ARAVA® was approved by the FDA on September 10, 1998, after the FDA Arthritis Advisory Committee unanimously recommended approval on August 7, 1998. Since 1998, ARAVA® has been used by over approximately 500,000 patients worldwide.

Proposed New Indication: Improvement in Physical Function

Currently, ARAVA® is indicated in adults in the U.S. for the treatment of active RA to reduce signs and symptoms of the disease and to retard structural joint damage as measured radiographically. The approval of ARAVA® was based on 1-year data from three Phase III, double-blind, randomized, controlled studies. Based on an analysis of 2-year data from these Phase III studies, it is proposed to add "and to improve physical function" to the currently approved indications.

These Phase III studies are summarized as follows:

- **US301** was a randomized, double-blind, placebo-controlled study of 482 patients, with a primary endpoint at 12 months and continued double-blind treatment to 24 months. Leflunomide was compared with both placebo and methotrexate (plus folate).

- **MN301/303/305.** Study MN301 was a randomized, double-blind, placebo-controlled, 6-month study of 358 patients, and the active comparator drug was sulfasalazine. Study MN303 was a double-blind, 6-month extension of MN301 without placebo control – patients who received placebo in MN301 were switched to sulfasalazine in a blinded manner at the start of MN303. Study MN305 was a double-blind extension of MN301/303 for a second year, during which patients who switched from placebo to sulfasalazine at the start of MN303 continued on sulfasalazine in MN305.

- **MN302/304.** Study MN302 was a 999-patient, randomized, double-blind study comparing leflunomide to methotrexate for 12 months. This study was not placebo-controlled, and concomitant folate administration was not required (only 10% of methotrexate patients received folate). Study MN304 was a double-blind extension of MN302 for a second year.

Patients in the Phase III, double-blind, randomized, controlled studies demonstrated that improvements observed at 6 and 12 months in ACR response criteria for signs and symptoms of RA, and in X-ray measurements of erosions and joint-space narrowing, were maintained over two years.

The Health Assessment Questionnaire (HAQ) was utilized to assess physical function in all three studies. In addition, the Problem Elicitation Technique (PET), 36-Item Short-Form (SF-36), Medical Outcomes Study (MOS) Current Health, and Work Productivity Questionnaire (WPQ) were used in one study (US301) as further measures of physical function and health-related quality of life.

At 12 and 24 months, leflunomide was statistically significantly superior to placebo in improving physical function and disability as assessed by the HAQ disability index (HAQ DI) and exceeded the generally accepted, 0.22-point change threshold for clinical significance. Superiority to placebo was demonstrated consistently across all eight HAQ DI subscales and the composite disability index in both placebo-controlled studies (US301 and MN301).

The SF-36 further addresses physical function as well as social and emotional function. In US301 at 12 months, leflunomide treatment resulted in statistically significant improvements compared to placebo in 5 of the 8 SF-36 scales (physical functioning, pain, general health perception, vitality, and social functioning), the SF-36 physical component summary score (PCS), and the Work Productivity Questionnaire (WPQ).

The improvements in physical function demonstrated at 6 and 12 months were maintained over 2 years. In patients continuing leflunomide for a second year of double-blind treatment, marked clinically meaningful improvements from baseline in the HAQ DI observed at month 6 and 12 were sustained at 24 months in all three trials, with no clinically meaningful changes between months 12 and 24. Improvements in the SF-36 observed at 6 and 12 months in Study US301 were maintained over 2 years.

A sensitivity analysis using three approaches to adjust for missing data within a treatment group showed superiority of leflunomide over placebo for the HAQ DI and SF-36 PCS. It thus demonstrated the robustness of the 2-year data.

Likewise, the adverse event profile of leflunomide during the second year of treatment was similar to that during the first year of treatment and no new types of adverse events emerged. The incidences of new-onset diarrhea, nausea, headache, alopecia, rash, hypertension, and increased liver function tests decreased in the second year of treatment.

The rates of adverse events seen in these phase III clinical trials with leflunomide were compared to methotrexate and sulfasalazine in a meta-analysis. The results show that serious and serious-and-related adverse events all occur more often amongst the methotrexate and sulfasalazine users.

In conclusion, these data and analyses support the efficacy of leflunomide with regard to improvement in physical function. Concomitant improvement in health-related quality of life was also demonstrated. Improvement in signs and symptoms was proven previously in 1-year studies contained in the NDA, and continued improvement over 2 years has been demonstrated in the extension studies.

Update on Safety of ARAVA®

The safety profile for ARAVA® is based on three types categories of safety information:

- Data from randomized, controlled, clinical trials
- Post-marketing safety surveillance data for ARAVA®
- Epidemiologic analysis of large cohorts of RA patients.

Comprehensive data on safety and adverse events from clinical trials and post-marketing surveillance for ARAVA® are reviewed in the Aventis response to a March 28, 2002 Public Citizen Health Research Group petition to the FDA requesting withdrawal of ARAVA® (leflunomide) Tablets from the market. The Aventis response was submitted to the FDA on August 8, 2002 and is included in *Appendix 1*. Rather than duplicate the review of clinical-trial and post-marketing surveillance data contained in *Appendix 1*, this *Briefing Document* will focus on the results of epidemiologic studies that assess the safety of leflunomide in comparison with other therapies for RA:

- **Retrospective cohort study.** The objective of this post-marketing, retrospective cohort study was to compare rates of adverse events (AEs) amongst leflunomide users to patients taking DMARDs (e.g., gold salts, azathioprine, hydroxychloroquine, D-penicillamine, sulfasalazine and the biologics etanercept and infliximab), and methotrexate (MTX), alone and in combination. This study relied on the 6.5 million-member claims database of Aetna, a US health insurer. Follow-up occurred from September 1998 through the end of December 2000. A diagnosis of Rheumatoid Arthritis and use of a DMARD were required for entry into the cohort. Subjects had to be 18 or over at time of entry. Exposure and time on drug was identified by dispensed prescription data. Outcomes included hepatic, hematologic, hypertensive, pancreatic, respiratory, and severe skin adverse events (AEs). Rates were computed using Poisson regression and were adjusted for age, sex, and comorbidities.

  The study assembled more than 40,500 RA patients and 83,000 person-years (PY) of follow-up, making it the largest RA cohort study ever conducted. The leflunomide monotherapy exposure group had significantly fewer AEs than DMARD and MTX groups. The leflunomide group had rates of hepatic, hematologic, pancreatic, pneumonitis, and severe skin AEs that were comparable to DMARD and MTX. Leflunomide patients had significantly lower rates of hypertension and upper respiratory AEs compared to DMARD and MTX. The combination of leflunomide + MTX exposure group had AE rates that were comparable to other combination therapies. The exposure group no-DMARDs generally had the highest rates observed in this study for all AEs. This is likely due to a 'depletion of susceptibles' effect and channeling bias, in which patients who experience an AE on a drug will be taken off and put on another, less toxic regimen.

  Although data on disease severity, OTC use, and history of RA were missing, it was clear that in this large population, leflunomide's safety profile is similar to that of other DMARDs.

- **Bi- cohort, nested, case-control study.** The objective of this study was to replicate the retrospective (Aetna) cohort study using different databases and a slightly different design. This study relied upon the combined data from the Protocare claims database (10 million members) and the PharMetrics database (16 million members). Follow-up occurred from September 1998 through December 2001. Subjects were entered if they had an RA diagnosis, had a prescription for a DMARD after September 1998, were 18 or over at entry, and had not experienced one of the endpoints of interest in the 90 days prior to entry. Exposures included methotrexate, leflunomide, other DMARDs, and biologic DMARDs.

The combined databases had almost 42,000 persons who were prescribed a DMARD after September 1998 and a total of 51,315 person-years of follow-up time. Three-quarters of the cohort were women. The average age of Protocare subjects was 59, compared to 49 for PharMetrics subjects. There were 90 events per 10,000 PY for all events of interest combined, and 5 per 10,000 PY for severe hepatic events, 27 per 10,000 PY for hematologic events, 16 per 10,000 PY for pancreatitis, 42 per 10,000 for opportunistic infections and sepsis, less than 1 per 10,000 PY for severe skin disease, 2 per 10,000 PY for pneumonitis, and 1 per 10,000 PY for lymphoma. Using methotrexate as the reference, the adjusted rate ratios for leflunomide were not significantly different from 1 for any serious adverse event (RR =1.1), serious hepatic events (RR = 0.9), serious hematologic events (RR = 0.8), serious pancreatitis events (RR = 1.5), and serious opportunistic infections and septicemia events (RR = 0.9). There were too few events for rate calculations of severe skin, pneumonitis, and lymphoma events. Of note were the generally elevated RRs for the biologic DMARDs, especially for any event, serious liver events, and opportunistic infections and septicemia events.

This study affirms the earlier Aetna cohort study in that adverse events amongst leflunomide patients do not occur more often than they do in methotrexate patients.

- **Proportional reporting ratio analysis.** The objective of the proportional reporting ratio (PRR) analysis was to determine if reports of adverse events amongst leflunomide users are inconsistent with similar reports amongst other DMARD users. PRR is a signal-generating tool, and is not used to confirm hypotheses. Proportional reporting ratio analysis compares spontaneous reports of suspected adverse reactions of different drugs where the true number of patients exposed to a drug is unknown.

  PRR analysis is widely employed by the Medicines Control Agency (MCA) in the UK. Criteria to evaluate the PRR come from several sources and are similar: a minimum of three reported cases are needed; a PRR of at least 3 and an associated X^2 over 5 or a PRR of at least 2 and an associated X^2 over 4 are considered possible signals. The data used are limited in that there is no way to assess the indication for a particular drug, so in the situation where a specific drug is used for more than one condition (e.g., as is the case with methotrexate), there is no way of adjusting for potential confounding by indication. The calculated PRR used the entire database of the FDA as a comparison (results which were not different than when DMARDs were used as a comparison group).

  The results showed that specific AE reports of leflunomide, as a proportion of all leflunomide reports was not different than other drugs, with the possible exceptions of interstitial lung disease PRR and vasculitis PRR. These signals have been further examined using epidemiologic data and have been found to be unsupported.

- **Reporting rate analyses.** The objective of this analysis was to examine the comparative reporting rates of various AEs of leflunomide and other DMARDs. Since this method relies on spontaneous report data, it is used for signal generation. Spontaneous reports (numerator data) were obtained from the FDA via QScan, a commercial software vendor that offers access to the more than two million adverse event cases reported to the FDA made available through the Freedom of Information Act, using proprietary mapping tools and techniques. Denominators (sales data) were obtained from IMS and converted into person-year exposures. Leflunomide, methotrexate, infliximab, and etanercept were compared. Adverse events of interest included hepatic failure, interstitial lung disease, tuberculosis and sepsis, bullous conditions, lymphoma, demyelinating disorders, hypertension, vasculitis, and pancytopenia.

Using this method to evaluate potential signals from spontaneous reports, none were found for leflunomide. Spontaneous report analysis is made difficult by under-reporting, the Weber effect (i.e., reports are more frequent closer to time of launch and for a period of about two years, then drop off substantially), lack of interest by professionals to report, potential confounding by indication (i.e., the AE is caused by the condition being treated, not the drug), and poor quality reporting data. Compared to the two biologic DMARDs, which were launched approximately the same time as leflunomide, there does not appear to be any signals.

Using this method of analysis, the AE profile of leflunomide appears comparable to that of biologic DMARDs, with lower rates for certain events.

- **Meta-analysis.** The objective of this study was to compare the rates of adverse events seen in phase III clinical trials; specifically, leflunomide was compared to methotrexate and to sulfasalazine. Adverse event rates were cumulated from clinical trials US301 (placebo-controlled trial of leflunomide versus methotrexate), MN301/303/305 (placebo-controlled trial and extensions of leflunomide versus sulfasalazine), and MN302/304 (leflunomide versus methotrexate). The rates are presented on a L'Abbé scatter plot (line-of-identity graph) for ease and sensibility of interpretation.

  The results of this meta-analysis show that Serious and Serious and Related adverse events all occur more often amongst the methotrexate and sulfasalazine users. Methotrexate and sulfasalazine also had higher rates of pain, blood, and cardiovascular AEs. Skin (rash) and hypertension occurred more often amongst leflunomide users. Leflunomide had higher rates of infection and abnormal liver tests compared to sulfasalazine, and lower rates compared to methotrexate.

  Using L'Abbé scatter plots to assess the rates of AEs reported in clinical trials of leflunomide, the two comparator agents (methotrexate and sulfasalazine) had higher rates of Serious and Serious and Related events, as well as higher rates of cardiovascular, blood, and pain AEs. Leflunomide had higher rates of skin rash and hypertension.

- **Liver transplant analysis.** The objective of this study was to determine how many liver transplant cases have been reported in which leflunomide or methotrexate is listed as the etiology. Data were requested by and received from the United Network for Organ Sharing (UNOS). UNOS administers the nation's only Organ Procurement and Transplantation Network (OPTN), established by the US Congress in 1984. Through the OPTN UNOS collects and manages data about every transplant event occurring in the United States; facilitates organ matching and placement processes; and helps to develop organ transplantation policy. All data on liver transplants from 1 January 1998 through 31 July 2002 were reviewed for drug involvement of either methotrexate or leflunomide.

  In 1998, three liver transplants listed methotrexate hepatotoxicity as a cause or diagnosis; in 1999 and 2000, one transplant each year listed methotrexate; in 2001, six transplants listed methotrexate; and through 31 July 2002, four cases listed methotrexate toxicity as a reason for the procedure. In that same time period, no cases listed leflunomide.

  Based on a review of the UNOS liver transplant data, methotrexate toxicity was listed as the diagnosis for 15 liver transplants from January 1998 through July 2002. In that same period, leflunomide toxicity was not listed as the diagnosis for liver transplant.

- **National Data Bank for Rheumatic Diseases.** Data from the National Data Bank for Rheumatic Diseases regarding rates of serious liver toxicity in patients taking leflunomide or methotrexate were published in abstract form and presented by Dr. Fred Wolfe at the American College of Rheumatology 2002 Annual Scientific Meeting. He reported that the rates were low and that there was no significant difference between leflunomide and methotrexate in the percent of patients with self-reported liver adverse events or in rates of liver adverse events per 100 patient-years [39].

  Treatment effectiveness in the community has been evaluated by Dr Fred Wolfe based on data from the National Data Bank for Rheumatic Diseases and reported at the American College of Rheumatology 2001 and 2002 Annual Scientific Meetings. Data were evaluated using the time patients remain on treatment [40] and also by an expanded definition of treatment failure, i.e., time to treatment discontinuation or addition of another DMARD [41]. In both of these measures of treatment effectiveness in the community, leflunomide and methotrexate were comparable.

Benefit-Risk Analysis of ARAVA®

The accepted standard of care for patients with RA is aggressive, early treatment with DMARDs, which slow and potentially alter the course of the disease. However, no single DMARD is effective in all patients, and secondary failures (loss of efficacy) are not uncommon. Accordingly, most patients with active RA require the progressive addition or change of treatments over time. Each of these therapies has been associated with serious and sometimes fatal adverse events, but this fact alone does not alter their positive benefit-risk profile. The need for alternative therapies remains the driving force behind recent development and approval of new treatments for RA over the last four years.

Several epidemiological approaches were used to compare the adverse event profile of leflunomide with those of other DMARDs. While none of the approaches are without limitations, the results of all analyses taken together show an adverse event profile for leflunomide comparable to methotrexate and other DMARDs.

The efficacy and safety data confirm that ARAVA® is an important advance in the treatment of RA and should remain available to the many thousands of individuals who benefit from the use of the drug. The chronic, progressive, and destructive nature of RA warrants the use of DMARDS early in the disease process. ARAVA® has been clinically proven to have efficacy in early and advanced disease, with rapid onset of therapeutic effect and sustained benefit during long-term therapy.

These established benefits must be weighed against its recognized risks, in the context of other available therapies and the severity of the disease. The risk of serious and sometimes fatal adverse events has, unfortunately, been observed with most prescription medications – and all DMARDs, including ARAVA®. Specifically, the safety data from randomized controlled trials show the overall percentage of patients with adverse events who were treated with ARAVA® was generally comparable to that of patients who received methotrexate and sulfasalazine. Importantly, nothing in the post-marketing experience changes the acceptable benefit-risk profile established by the controlled clinical studies.

When weighed against the benefits of the drug, its impact on the disease course, and the limitations of other available therapies, the risks of ARAVA® treatment are clearly outweighed by its substantial benefits.

LIST OF ABBREVIATIONS

ACR	American College of Rheumatology
ALT	alanine transaminase
ANCOVA	analysis of covariance
AST	aspartate transaminase
CI	confidence interval
DHO-DH	dihydroorotate dehydrogenase
DMARD	disease-modifying antirheumatic drug
HAQ	Health Assessment Questionnaire
HAQ DI	Health Assessment Questionnaire Disability Index
HWA 486	leflunomide
ITT	intention-to-treat
LEF	leflunomide
LOCF	last observation carried forward
MACTAR	McMaster Toronto Arthritis Patient Preference Disability Questionnaire
MCID	minimum clinically important difference
MCS	mental component summary score
MN301	study HWA486/6/MN/301/RA
MN302	study HWA486/6/MN/302/RA
MN303	study HWA486/6/MN/303/RA
MN304	study HWA486/6/MN/304/RA
MN305	study HWA486/6/MN/305/RA
MOS	Medical Outcomes Study
MTX	methotrexate
NSAID	nonsteroidal anti-inflammatory drug
OMERACT	outcome measures in rheumatoid arthritis clinical trials
PCS	physical component summary score
PET	Problem Elicitation Technique
PBO	placebo
RA	rheumatoid arthritis
SD	standard deviation
SF-36	36-Item Short-Form
SSZ	sulfasalazine
ULN	upper limit of normal range
US301	study HWA486/F/US/301/RA
WPQ	Work Productivity Questionnaire

1. IMPROVEMENT IN PHYSICAL FUNCTION

1.1 Background Information

1.1.1 General background

Leflunomide is a pyrimidine synthesis inhibitor that acts as an antiproliferative agent. Following oral administration it is rapidly metabolized to an active metabolite A77 1726 (hereafter referred to as M1). M1 has been shown to be active in vitro and is presumed to be the active metabolite in vivo. In vitro, M1 inhibits mitogen-stimulated proliferation of human peripheral blood mononuclear cells and transformed murine and human cell lines in a dose-dependent fashion. This antiproliferative activity is reversed by the addition of uridine to the cell culture, indicating that M1 acts at the level of the pyrimidine biosynthesis pathway. Binding studies using radiolabeled ligand demonstrate that the active metabolite binds to the human enzyme dihydroorotate dehydrogenase (DHO-DH). In vitro incubation of M1 with rat, mouse, and human DHO-DH demonstrated that it inhibited the activity of the enzyme at concentrations lower than those that exert antiproliferative effects upon rapidly dividing cells: 16–657 nM. Rat and mouse enzymes are more sensitive to this inhibitory effect (IC_{50} 0.14±0.08 and 16±11 μM, respectively) than the human enzyme (IC_{50} 46±6 μM). Together, these data suggest that at concentrations achievable in subjects *de novo* pyrimidine synthesis in activated lymphocytes and other rapidly dividing cell populations is inhibited, resulting in reversible cell cycle arrest.

ARAVA® (leflunomide) received NDA approval on 10 September 1998 for the treatment of active RA to reduce signs and symptoms and to retard structural damage as evidenced by erosions and joint space narrowing shown in radiographs of the hands and feet. The FDA requires 2-year data before considering a claim based on improvement in physical function; however, only 1-year physical function and health-related quality of life data were available at the time of approval (summarized in this document in *Section 1.1.3.2 Year-1 results in ITT population)*. The Sponsor has therefore submitted a supplemental NDA providing 2-year data from three Phase III studies in support of an indication for improved physical function (summarized in *Section 1.1.3.3 Results for Year-2 Cohort*). Evidence of improvement in physical function has been based on the Health Assessment Questionnaire (HAQ), Problem Elicitation Technique (PET), and Medical Outcomes Study (MOS) 36 item short form survey (SF-36), which are described in *Section 1.1.2.1 Scales used for patient-reported outcome measures*.

1.1.2 Overview of clinical studies

The following is a background and overview of the clinical investigations presented in the supplemental NDA. Summaries of the efficacy and safety results from these studies are provided in *Sections 1.3 to 1.5* of this *Briefing Document*.

This *Briefing Document* details the findings from three long-term (up to 24 months of treatment), Phase III trials supporting a claim for improved physical function. All three of the trials were multinational, multicenter, double-blind, parallel-group studies and extended to 2 years. The approval of ARAVA® was based on 1-year data from the three Phase III pivotal trials that were submitted in the original NDA. One of these studies, HWA486/F/USA/301/RA (US301), compared

leflunomide with methotrexate over 2 years and compared both of these medications with placebo. A second study, HWA486/6/MN/304/RA (MN304), compared leflunomide with methotrexate in the second year of active treatment. The third study, HWA486/6/MN/305/RA (MN305), compared leflunomide with sulfasalazine in a second year of active treatment. All three studies gathered information on patient physical function. US301 also collected data on general health-related quality of life for 2 years. Following are brief descriptions of each of these three studies.

1.1.2.1 US301

US301 was designed as a 2-year double-blind trial to provide long-term data on the safety and efficacy of leflunomide compared with placebo and methotrexate in the treatment of patients with active RA. The primary, protocol-defined endpoint for the study was the ACR20 response after 12 months of treatment in the initial therapy phase. This primary 12-month analysis for patients treated at the 42 US study sites was previously reported in the original NDA submission. The study synopsis for this study provided in *Appendix 2* presents results for the full 24 months of the initial therapy phase as well as results from the alternate therapy phase (1-year data) for patients treated at the 42 US and 5 Canadian study sites. A summary of the study design and patient disposition is given in the Figure on page 13.

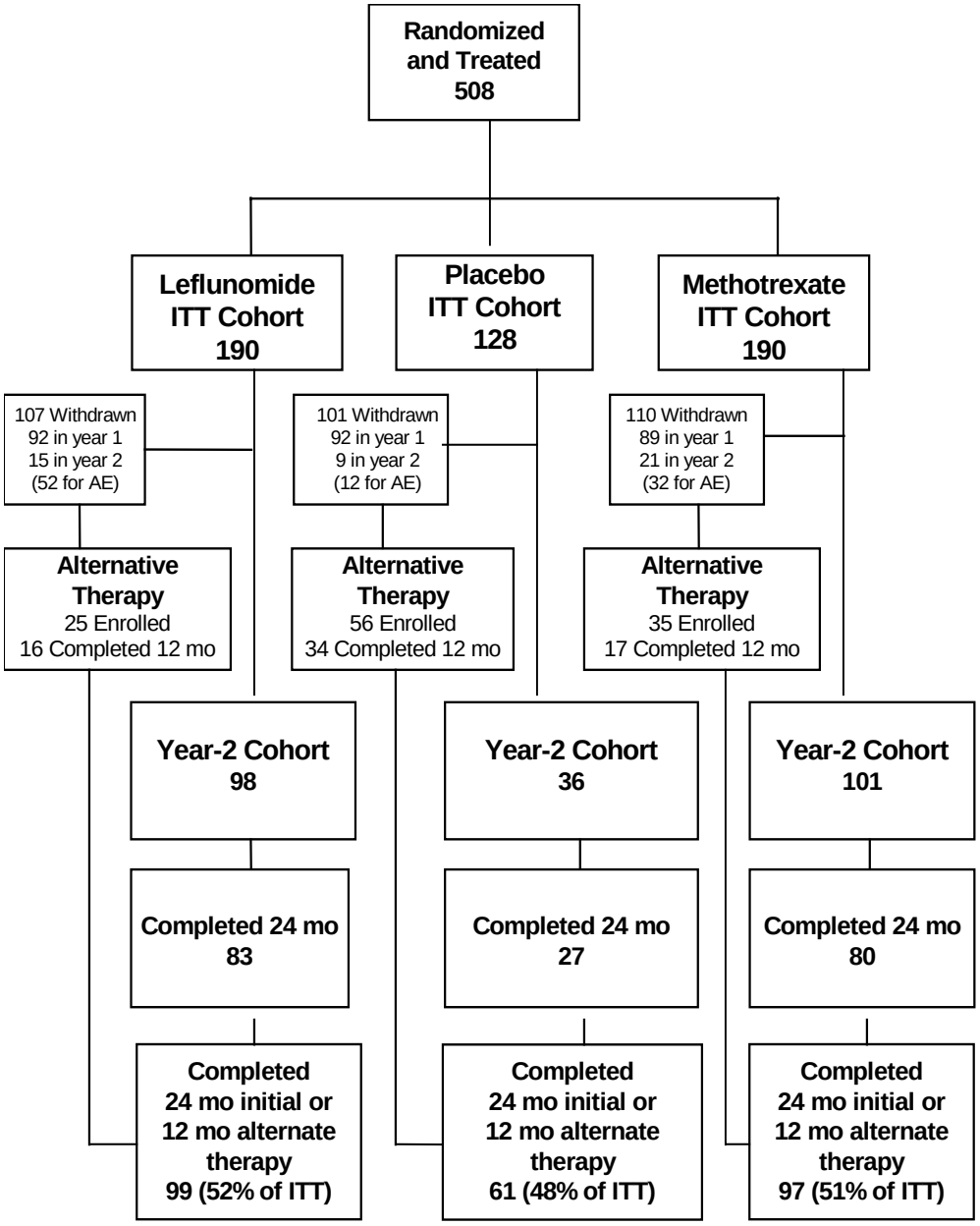
Patients were eligible for the alternate therapy phase if they had received at least 16 weeks of study treatment and were withdrawn due to lack of clinical response, significant toxicity, or persistent laboratory abnormalities. Patients who were randomized to methotrexate or placebo in the initial therapy phase received leflunomide in the alternate therapy phase; patients originally randomized to receive leflunomide received methotrexate.

In support of a claim for improved physical function, the following questionnaires were used in US301 to gather information on patients' physical function and health-related quality of life at baseline, weeks 24, 52, 104, or when the patient switched medication or left the study.

- Health Assessment Questionnaire (HAQ)
- Problem Elicitation Technique (PET)
- 36-Item Short-Form (SF-36)
- Medical Outcomes Study (MOS) current health perceptions scale
- Work Productivity Questionnaire (WPQ): Items on work problems and work productivity were abstracted from the 1994 National Opinion Research Center Survey [11].

Patient accountability in US301 is summarized in the following tables. A small number of Canadian patients accrued late due to lack of drug supply were not analyzed in the original NDA submission but only in the 2 year safety and efficacy data. 'Alternate therapy' was offered to all patients who exited initial therapy due to documented lack of efficacy on or after 4 months.

Results for the alternate therapy phase are not part of the analyses presented in this document, but are included in the accountability data as patients remained in protocol treatment.



Patient Disposition for Study US301

Patient accountability in US301

	No. (%) subjects			
Patient status	**Leflunomide**	**Placebo**	**Methotrexate**	**Total**
Enrolled (ITT cohort)	190 (100%) 8 from Canad	128 (100%) 10 from Canada	190 (100%) 8 from Canada	508 (100%)
Entered 2nd year of treatment	98 (52%)	36 (28%)	101 (53%)	235 (46%)
Completed 24 months of treatment	83 (44%) 85% of those entering 2nd yr	27 (21%)	80 (42%) 79% of those entering 2nd yr	190 (37%)
Alternate therapy*				
Enrolled	25 to MTX 1 from Canad	56 to LEF 5 from Canada	35 to LEF 2 from Canada	116
Completed 1 year on new therapy	16	34	17	67
Completed entire blinded treatment (24 months on initial therapy or 12 months on alternate therapy)	99 (52%)	61 (48%)	97 (51%)	257 (51%)

* Results for the alternate therapy phase are not included in the analyses presented in this document.

Patient accountability for one year data in US301 as submitted in original NDA
[includes ONLY patients enrolled in US]

	No. (%) subjects			
Patient status	**Leflunomide**	**Placebo**	**Methotrexate**	**Total**
Enrolled (ITT cohort)	182 (100%)	118 (100%)	182 (100%)	482 (100%)
Completed 12 months treatment On initial therapy	96 (53%)	37 (31%)	105 (58%)	
Alternate therapy*				
Eligible	30 (16%)	60 (51%)	42 (23%)	
Enrolled	24 to MTX	51 to LEF	33 to LEF	108
Completed 12 months treatment in initial &/or alternate therapy	120 (66%)	88 (75%)	138 (77%)	346 (72%)

1.1.2.2 MN305

MN305 was a double-blind extension to the 24-week study MN303, which in turn was a double-blind extension to the 24-week study MN301. The objectives of the studies MN301/303/305 were to investigate the safety of leflunomide during long-term use in RA patients, to assess the relative efficacy and safety profile of leflunomide compared with sulfasalazine, and to investigate population pharmacokinetics. In the original protocol for MN305, blinded treatment was to be continued until

the database for MN301 had been unblinded. The double-blind treatment period was, however, subsequently extended in two amendments to allow patients to complete 2 years of treatment. The amendments addressed recommendations made by the FDA, European regulatory authorities, and independent experts in the field of rheumatology. Patients who received leflunomide or sulfasalazine in MN301 continued on their respective medication in MN303 and MN305. Patients who received placebo in MN301 were switched to sulfasalazine in a blinded manner at the start of MN303 and continued on sulfasalazine in MN305 (placebo/sulfasalazine group). At the start of MN305, all patients continued on the same daily dosage of leflunomide or sulfasalazine that they had been taking at the completion of MN303. Results of MN301 and MN303 were previously reported in the original NDA submission.

The HAQ was used in MN301/303/305 to gather information on patient functional impairment for 2 years. Assessments were made at baseline, weeks 2, 4, 6, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 72, 104, or when the patient discontinued early.

Patient accountability in MN301/303/305 is summarized in the following table:

Patient accountability in MN301/303/305

	No. (%) subjects			
Study/patient status	**Leflunomide**	**Placebo**	**Sulfasalazine**	**Total**
MN301				
Enrolled (ITT cohort)	133 (100%)	92 (100%)	133 (100%)	358 (100%)
Completed (6 months)	96 (72%)	51 (55%)	83 (62%)	230 (64%)
MN303				
Enrolled*	80 (60%)	41 to SSZ**	76 (57%)	197 (55%)
Completed (12 months)	71 (53%)	29	68 (51%)	168 (47%)
MN305				
Enrolled*	60 (45%)	26	60 (45%)	146 (41%)
Completed (24 months)	53 (40%)	21	47 (35%)	121 (34%)

* Some patients who completed MN301 or MN303 elected not to continue in the next extension protocol. These patients were equally distributed between responders and nonresponders.

** Patients who received placebo in MN301 were switched to sulfasalazine in a blinded manner at the start of MN303 and continued on sulfasalazine in MN305. These patients are not included in the analyses of the sulfasalazine treatment group.

1.1.2.3 MN304

MN304 was a double-blind extension of study MN302 to investigate the safety of leflunomide during long-term use in RA patients, to assess the relative efficacy and safety profile of leflunomide compared with methotrexate during long-term treatment, and to investigate population pharmacokinetics. In the original protocol for MN304, blinded treatment was to be continued until the database for MN302 had been unblinded. The double-blind treatment period was, however, subsequently extended by an amendment to allow all patients to complete 2 years of treatment. The amendment addressed recommendations made by the FDA, European regulatory authorities, and independent experts in the field of rheumatology. At the start of MN304, all patients continued on the same dosage of leflunomide or methotrexate that they had been taking at the completion of MN302. Results of MN302 were previously reported in the original NDA submission.

The HAQ was used in MN302/304 to gather information on patient functional impairment over 2 years. Assessments were made at baseline, weeks 2, 4, 6, 8, 12, 24, 36, 52, 76, 104, or when the patient discontinued early.

Patient accountability in MN302/304 is summarized in the following table:

Patient accountability in MN302/304

	No. (%) subjects		
Study/patient status	**Leflunomide**	**Methotrexate**	**Total**
MN302			
Enrolled (ITT cohort)	501 (100%)	498 (100%)	999 (100%)
Completed (12 months)	349 (70%)	387 (78%)	736 (74%)
MN304			
Enrolled*	292 (58%)	320 (64%)	612 (61%)
Completed (24 months)	256 (51%)	277 (56%)	533 (53%)

* Some patients who completed MN302 elected not to continue in MN304. These patients were equally distributed between responders and nonresponders.

1.2 Physical Function/Health-Related Quality of Life Instruments

Physical function / health-related quality of life were assessed in US301 by means of the HAQ, PET, SF-36, MOS Current Health, and WPQ. The HAQ was used in MN304 and MN305. This section describes each of these measures used for the statistical analyses that are presented in the supplemental NDA. Physical function and health-related quality of life instruments are often referred to as "patient-reported outcome" measures. This collective term will be used in the subsequent sections of this *Briefing Document*.

1.2.1 Scales used for patient-reported outcome measures

The following table summarizes the physical function/health-related quality of life characteristics that each scale measures. Copies of patient questionnaires for these scales are provided in *Appendix 3*.

Physical function / health-related quality of life assessment scales

Scale	Item number	Focus of scale
HAQ		
Overall difficulty and weighted difficulty	1-24	difficulty with performing basic tasks
Overall health	25	perception of overall health
PET	1-24	difficulty and importance of performing basic tasks
SF-36		
Physical functioning	3a-3j	limitations of physical function
Role limitations due to physical problems	4a-4d	difficulty performing usual activities due to physical problems
Bodily pain	7, 8	amount of discomfort and its interference with activities
General health	1, 11a-d	perception of overall personal health
Vitality	9a, 9e, 9g, 9i	pep and energy
Social functioning	6, 10a	social contacts and activities
Role limitations due to emotional health	5a-5c	difficulty performing usual activities due to emotional problems
Mental health	9b, 9c, 9d, 9f, 9h	depression and anxiety
Health transition	2	comparison of health to a year ago
MOS Current Health		
Current health perceptions	11e-f	perception of current health
WPQ		
Work problems	14 a-d	frequency of work problems
Work productivity	15 a-j	difficulty performing work tasks

1.2.1.1 Health Assessment Questionnaire (HAQ)

The HAQ is a validated instrument developed to assess disease-specific physical function and degree of disability in patients suffering from RA [5, 13]. It consists of various questions relating to eight categories (dressing and grooming, rising, eating, walking, hygiene, reach, grip, and activities). HAQ is one of the two instruments that are recognized by the FDA in the 1999 Guidance document for assessing the prevention of disability [4].

The HAQ Disability Index (HAQ DI) uses the scores of the worst items within each of the eight categories, modified by the use of devices and aids. In the event that a patient indicates that he/she uses a device or aid to perform a task, the score for the associated category increases to 2 ("able to do with much difficulty") if it was previously 0 or 1.

1.2.1.2 Problem Elicitation Techniques (PET)

The self-completed PET questionnaire was derived from the interview-based McMaster Toronto Arthritis Patient Preference Disability Questionnaire (MACTAR). The HAQ investigates the level of difficulty patients report when performing each item in the questionnaire. The PET builds on the HAQ by inquiring which of these physical activities are most affected by RA and which they would most like to see improved by treatment. Patients rank the difficulty, severity, and/or frequency of performing these activities and are then asked to rate their level of importance. Therefore, the PET is structured to demonstrate changes in physical function judged important to the patient.

The weighted top 5 score of the PET is calculated as follows: the difficulty score for each patient is first multiplied by its importance, all items are then ranked and the top 5 items are averaged to give the weighted top 5 score.

1.2.1.3 Medical Outcomes Study 36-Item Short Form (SF-36)

The SF-36 is a generic health-related quality of life instrument and is also recommended in the 1999 FDA Guidance, in addition to being a disease-specific instrument [4]. It has proved to be valid and reliable in a large number of indications and patient populations (e.g., cardiovascular disease, low back pain, diabetic foot ulcers, total knee replacement, and dialysis) [26]. It was developed in the US and designed to represent eight of the most important health concepts: physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health. In addition, a single question assesses reported health transition. Each domain generates a transformed score in the range 0–100, with 0 being the worst score and 100 the best.

In addition to the scores for the eight SF-36 domains, two summary scores were calculated for the physical and mental component summary scores (PCS and MCS, respectively) as described in the SF-36 Physical and Mental Health Summary Scales User's Manual [25]. The two summary scores were calculated based on a weighted linear combination of the eight SF-36 domains. The instrument was developed so that the general population has a mean score of 50 with a standard deviation of 10.

1.2.1.4 MOS Current Health

The MOS Current Health scale is a generic measure of health status. It was developed by the RAND Corporation to assess the subject's own rating of overall current health [18]. The MOS Current Health scale generates a transformed score in the range 0–100, with 0 being the worst score and 100 being the best.

1.2.1.5 Work Productivity Questionnaire (WPQ)

The WPQ comprises 14 questions that measure "on the job" impact of RA and its treatment [11]. This self-assessment instrument was used to measure the degree to which chronic health problems interfered with the ability to perform job roles.

The WPQ was used because the SF-36 scales examining disability are relatively broad and only distinguish a limited range of disability levels. The WPQ was designed to fill this gap and measure "on the job" impact of RA and its treatment.

1.2.2 Statistical methodology for patient-reported outcomes

The 2-year data on physical function and health-related quality of life were obtained from the three pivotal studies and analyzed in support of this claim. The report for the 2-year data, from which information for this *Briefing Document* was extracted, is provided in the supplemental NDA.

The primary endpoint for efficacy was at 12 months in US301 and MN302 and at 6 months in MN301. Treatment group comparisons of leflunomide versus placebo and active controls (methotrexate or sulfasalazine) were conducted at these primary endpoints. Maintenance of effects was evaluated between 12 and 24 months of treatment within treatment groups.

The analyses of patient-reported outcomes presented in this document were performed on the intention-to-treat (ITT) and year-2 cohorts in order to evaluate physical function and health-related quality of life:

- The **ITT cohort** includes all patients initially enrolled into US301, MN301, and MN302 who took at least one dose of study medication.
- The **year-2 cohort** includes all patients initially enrolled into US301, MN301, and MN302 and who entered a 2nd year of treatment.

The following comparisons were performed:

- Leflunomide versus placebo in the ITT population using last-observation-carried-forward (LOCF) methodology at the primary endpoint for each protocol and at 24 months
- Leflunomide versus methotrexate in the ITT population using LOCF at the primary endpoint for each protocol and at 24 months
- Leflunomide versus sulfasalazine in the ITT population using LOCF at the primary endpoint for each protocol and at 24 months
- Leflunomide versus methotrexate in the year-2 cohort using LOCF
- Leflunomide versus sulfasalazine in the year-2 cohort using LOCF
- To assess the maintenance of effects during the second year of treatment, the 24-month data were compared with the 12-month data within each treatment group for the year-2 cohort using LOCF.

In the year-2 cohort, statistical comparisons with placebo were not performed due to the predictably small number of placebo patients.

Treatment comparisons for the ITT population were performed by analysis of covariance (ANCOVA) with treatment, region, disease duration, time since last disease-modifying antirheumatic drug (DMARD), pairwise interactions with treatment as factors, and baseline as covariate. For the analysis of the year-2 cohort, baseline imbalances for the above-mentioned covariates were not significant. Therefore the model was reduced to treatment, region, and treatment x region interaction.

In addition, a sensitivity analysis was performed to evaluate the robustness of the analyses described above. The methodology for the sensitivity analysis is described in *Section 1.1.3.4.2.1 Methodology of sensitivity analysis*.

1.3 Results for Patient-Reported Outcomes

1.3.1 Rationale for presentation of patient-reported outcomes

Placebo-controlled clinical trials are still considered to provide the most convincing basis for assessing the efficacy and safety of a compound. However, in RA patients, placebo-controlled trials (especially long-term trials of at least 2 years' duration) are difficult to conduct and in the future will be increasingly difficult to justify [27].

In view of the fact that it is difficult to conduct long-term placebo-controlled studies of at least 2 years, all available evidence on efficacy should be used as a basis for an indication claim. In addition, the high withdrawal rate in the placebo group in study US301 makes it even more difficult to base the assessment of efficacy solely on an ITT analysis. The following rationale for the presentation of the results was used to allow a comprehensive assessment of the effects of leflunomide on patient-reported outcomes:

- ITT analysis of all studies and all treatment groups after 1 year of treatment (*Section 1.1.3.2 Year-1 results in ITT population*)
- analysis of year-2 cohort including maintenance of treatment effects within the active treatment groups (*Section 1.1.3.3 Results for Year-2 Cohort*)
- ITT analysis of all studies for the 2-year data (including the comparison of leflunomide with placebo in study US301) (*Section 1.1.3.4.1 LOCF analysis*)
- a supportive sensitivity analysis to assess the robustness of the ITT 2-year analysis with regard to the missing data caused by early withdrawals (*Section 1.1.3.4.2 Sensitivity analysis*).

The evidence from these approaches forms the basis for demonstrating the efficacy of leflunomide in improving physical function in RA patients.

1.3.2 Year-1 results in ITT population

The ITT cohort includes all patients initially enrolled into US301, MN301, and MN302 who took at least one dose of study medication.

1.3.2.1 HAQ DI results at 6 and 12 months (US301, MN301/303 and MN302)

The following table summarizes the mean changes in the HAQ DI for each Phase III study.

Mean change from baseline to endpoint for the HAQ DI (ITT population)

Study/treatment group	N	Baseline HAQ DI	Mean change (%) at endpoint
US301 (12 months)			
Leflunomide	166	1.30	–0.45*, ** (–35%)
Placebo	101	1.31	–0.03 (–2%)
Methotrexate	169	1.30	–0.26 (–20%)
MN301 (6 months)			
Leflunomide	113	1.65	–0.56*, *** (–34%)
Placebo	81	1.59	–0.08 (–5%)
Sulfasalazine	111	1.50	–0.37 (–25%)
MN301/303 (12 months)			
Leflunomide	65	1.68	–0.67 (–40%)
Sulfasalazine	61	1.42	–0.53 (–37%)
MN302 (12 months)			
Leflunomide	464	1.50	–0.44*** (–29%)
Methotrexate	463	1.52	–0.54 (–36%)

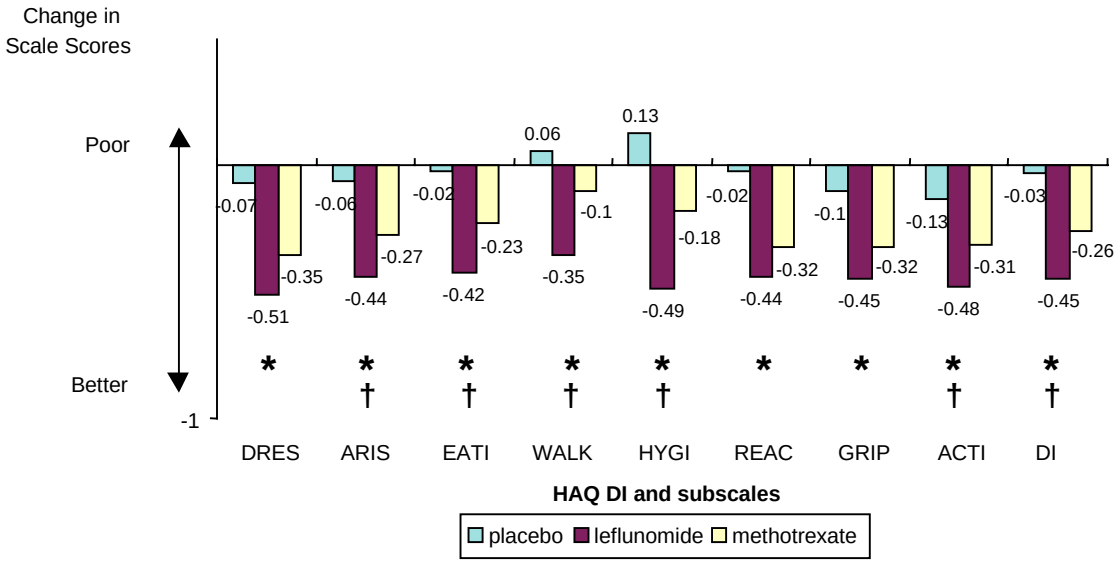
* Indicates statistically significant difference between leflunomide and placebo ($p \leq 0.001$).
** Indicates statistically significant difference between leflunomide and active comparator ($p \leq 0.01$).
*** Indicates statistically significant difference between leflunomide and active comparator ($p \leq 0.05$).

Leflunomide treatment demonstrated statistically significantly greater improvements in physical function than placebo. In US301 and MN301 statistically significantly greater improvement occurred with leflunomide than with methotrexate after 12 months or with sulfasalazine after 6 months. Conversely, in MN302 the methotrexate group showed statistically significantly greater improvement at 12 months, although notable improvement also occurred in the leflunomide group.

It has been suggested that improvements of 36% from baseline values or 18% better than placebo would be clinically important in RA patients [6]. More specifically, it has been noted that a change of –0.22 points in the HAQ DI (i.e., an improvement of 0.22) reflects a clinically meaningful change [27]. In the three Phase III studies, the results of the leflunomide treatment groups met or exceeded this level: the mean HAQ DI improved by 0.44 to 0.67, reflecting clinically meaningful changes.

Changes between baseline and month 12 for the HAQ DI and its subscales in US301 are shown in the following diagram.

Study US301: ITT population
Mean changes from baseline to month 12 in HAQ DI and its subscales



* Leflunomide is significantly better than placebo at 0.05 level of significance.
† Leflunomide is significantly better than methotrexate at 0.05 level of significance.
Note: Number of patients varies between scales

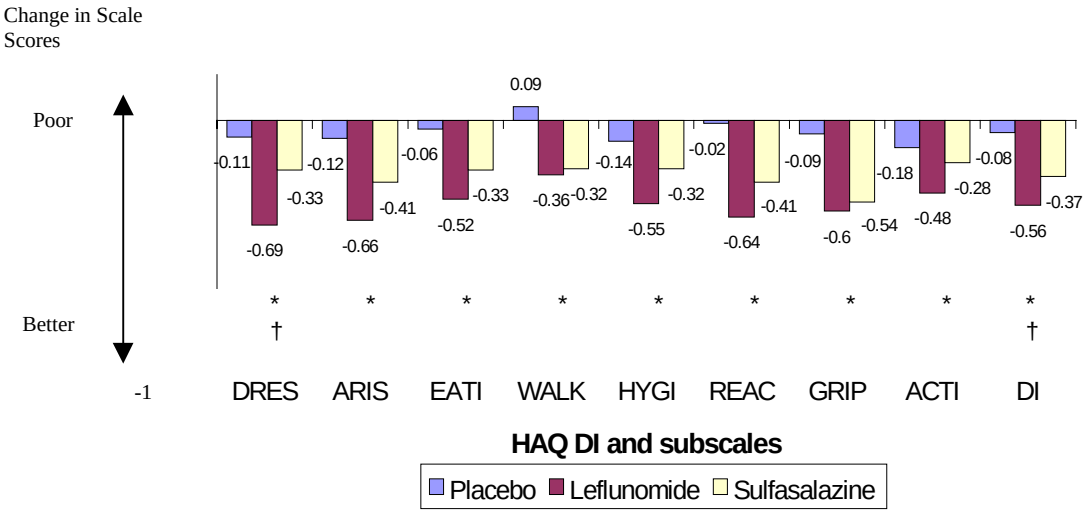
Scale abbreviations:

DRES	= Dressing	WALK	= Walking	GRIP	= Grip
ARIS	= Arising	HYGI	= Hygiene	ACTI	= Activities
EATI	= Eating	REAC	= Reach	DI	= Disability Index

In US301, statistically significant improvements from baseline to endpoint in the HAQ DI and in all 8 subscale scores were evident in the leflunomide treatment group compared to placebo and 5 of the 8 subscales compared to methotrexate. The magnitude of improvement in all subscales in the leflunomide-treated group was clinically meaningful (>0.22) and approached 0.5 in most of the subscales [7,27].

Changes between baseline and month 6 for the HAQ DI and its subscales in **MN301** are shown in the following diagram.

Study MN301: ITT population
Mean changes from baseline to month 6 in HAQ DI and its subscales



* Leflunomide is statistically significantly better than placebo at 0.01 level of significance.
† Leflunomide is statistically significantly better than sulfasalazine at 0.05 level of significance.

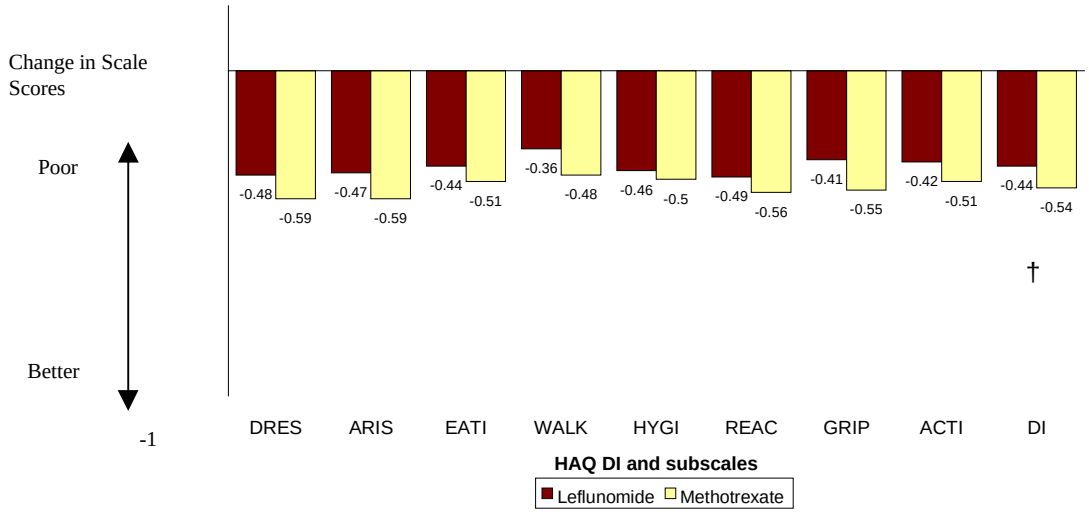
Scale abbreviations:

DRES = Dressing	WALK = Walking	GRIP = Grip
ARIS = Arising	HYGI = Hygiene	ACTI = Activities
EATI = Eating	REAC = Reach	DI = Disability Index

In MN301, statistically significant improvements from baseline to endpoint in the HAQ DI and in all HAQ subscale scores were evident in the leflunomide treatment group compared to placebo and in the HAQ DI and 1 of the 8 subscales compared to sulfasalazine after only 6 months of therapy. Again, in the leflunomide group these improvements correspond to score changes exceeding –0.22 points and are clinically meaningful.

Changes between baseline and month 12 for the HAQ DI and its subscales in MN302 are shown in the following diagram.

Study MN302: ITT population
Mean changes from baseline to month 12 in HAQ DI and its subscales



† Leflunomide is statistically significantly poorer than methotrexate at 0.05 level of significance.

Scale abbreviations:

DRES = Dressing	WALK = Walking	GRIP = Grip
ARIS = Arising	HYGI = Hygiene	ACTI = Activities
EATI = Eating	REAC = Reach	DI = Disability Index

In MN302, improvement in the leflunomide-treated patients was statistically significantly lower than in methotrexate patients for the HAQ DI, but not for the individual subscales. Nonetheless, improvements in the leflunomide group were clinically meaningful and approached –0.5 in each subscale. They were of similar magnitude to the improvements observed in US301 and MN301 and demonstrate a consistent treatment effect across the three phase III trials.

The difference between the HAQ DI in the leflunomide and methotrexate groups was very small and considerably less than the minimum clinically important difference for HAQ DI of 0.22 [27]. Thus, although the observed difference between the treatment groups was statistically significant due to the large sample sizes, it may not be clinically meaningful.

1.3.2.2 PET results for year 1 (US301 only)

The PET weighted top 5 score ranks the five activities most important to the patient [1,2]. The results of the weighted top 5 score in US301 are summarized in the table below.

Study US301: ITT population
Mean changes from baseline to month 12 in PET weighted top 5 score

Parameter	Leflunomide	Placebo	Methotrexate
No. of patients	166	101	170
Baseline mean	21.2	22.4	20.4
Mean change	–6.91**, ***	–0.66	–3.41
Mean % change	35%	3%	17%
Mean % improvement vs. placebo*	32%	--	13%

* Percent improvement versus placebo was calculated as follows:
(LEF mean change from baseline – Placebo mean change from baseline)/LEF baseline mean

** Indicates statistically significant difference between leflunomide and placebo ($p \leq 0.0001$).

*** Indicates statistically significant difference between leflunomide and methotrexate ($p \leq 0.01$).

Results in the leflunomide group were statistically superior to placebo and to methotrexate. The PET results are, in a sense, the most sensitive measure of physical function in that the data are "customized" for each patient and thus reflect improvements in the performance of those activities most important to each patient.

Mean changes in PET in US301 replicate the changes in HAQ DI. These combined data indicate improvement in performance of physical activities important to patients and, in particular, those activities in leflunomide-treated subjects.

The frequencies of the ten categories that were most often selected by each of the patients as his or her "top 5" are listed in the following table.

Study US301: ITT population
Frequencies of the ten most commonly selected PET top 5 categories

Category	No. patients	% patients
Do chores	204	42.5
Stand from chair	203	42.3
Dressing self	195	40.6
Get in/out of bed	163	34.0
Get down 5-lb bag	160	33.3
Open milk carton	148	30.8
Take a tub bath	147	30.6
Open jars previously opened	145	30.2
Shampoo hair	118	24.6
Climb up 5 steps	112	23.3

1.3.2.3 SF-36, MOS Current Health, and WPQ results for year 1 (US301 only)

Changes in the SF-36 PCS, SF-36 MCS, MOS Current Health, and WPQ in US301 are summarized in the table below:

Study US301: ITT population
Mean changes from baseline to month 12 for SF-36 summary scores, MOS Current Health, and WPQ

Parameter	Leflunomide	Placebo	Methotrexate
SF-36 PCS			
No. of patients	157	101	162
Baseline mean	30.0	28.9	29.7
Mean change	7.6**, ***	1.0	4.6
Mean % change	25%	3%	15%
Mean % improvement vs. placebo*	22%	--	12%
SF-36 MCS			
No. of patients	157	101	162
Baseline mean	46.8	48.3	48.5
Mean change	1.5	0.8	0.9
Mean % change	3%	2%	2%
Mean % improvement vs. placebo*	1%	--	0%
MOS Current Health			
No. of patients	156	100	164
Baseline	44.8	41.1	42.0
Mean change	8.7	4.2	9.8
Mean % change	19%	10%	23%
Mean % improvement vs. placebo*	9%	--	13%
Work Productivity Questionnaire			
No. of patients	138	92	148
Baseline mean	53.3	52.9	51.9
Mean change	9.8**	0.3	7.5
Mean % change	18%	0.5%	14%
Mean % improvement vs. placebo*	18%	--	14%

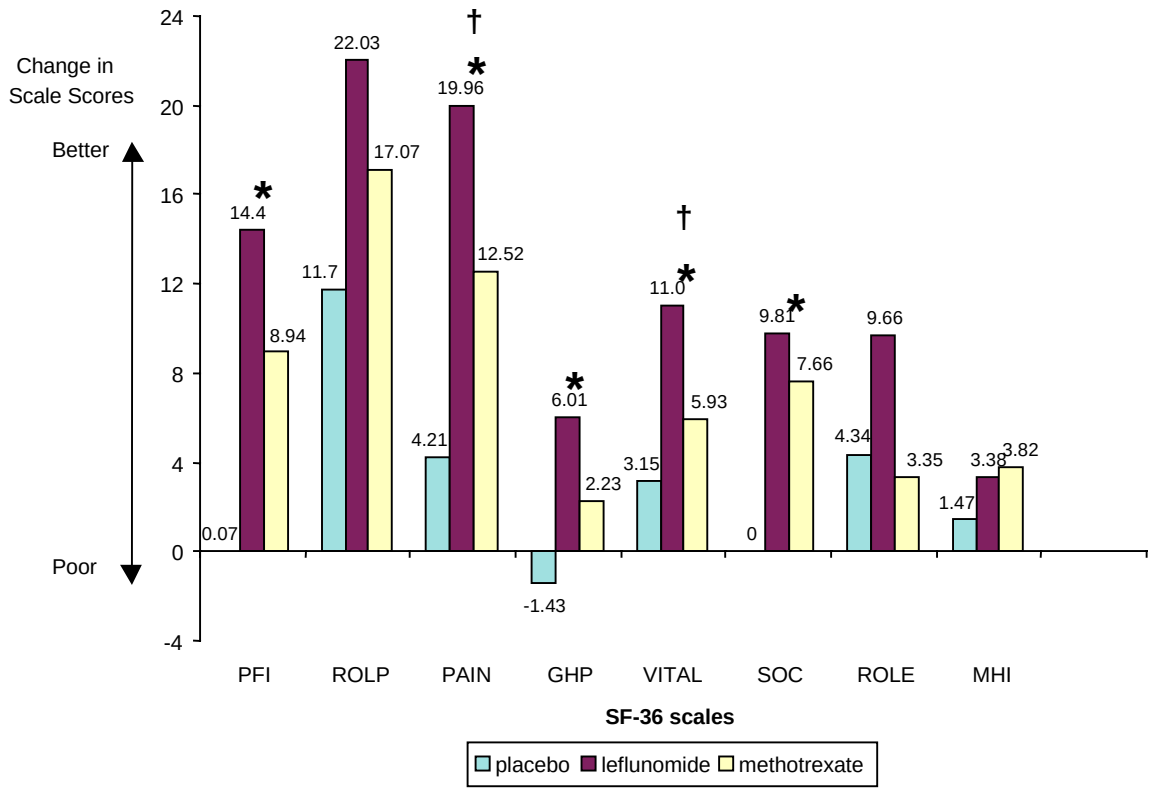
* Percent improvement versus placebo was calculated as follows:
(LEF mean change from baseline – Placebo mean change from baseline)/LEF baseline mean
** Indicates statistically significant differences between leflunomide and placebo p <0.001.
*** Indicates statistically significant difference between leflunomide and methotrexate p <0.02.

There was a statistically significant difference between the leflunomide and placebo groups in the SF-36 PCS, but not in the SF-36 MCS. The MCS improved only slightly, but this was expected because patients had baseline scores close to normal. The MOS Current Health score did not demonstrate a statistically significant difference between treatment groups.

Mean changes in work productivity showed statistically significant improvement with leflunomide treatment compared to placebo (p=0.0024) but not compared to the methotrexate treatment group. This is interpreted to mean that patients treated with leflunomide had higher productivity at work, home, or school.

The figure below summarizes mean changes in the subscale scores of the SF-36 and demonstrates the efficacy of leflunomide compared to placebo and methotrexate, as judged by the patient.

Study US301: ITT population
Mean changes from baseline to month 12 in SF-36 scales



* Leflunomide is significantly better than placebo at 0.05 level of significance.
† Leflunomide is significantly better than methotrexate at 0.05 level of significance.
Note: Number of patients varies between scales

Scale abbreviations:

PFI = Physical Functioning	GHP = General Health Perception	ROLE = Role Emotional
ROLP = Role Physical	VITAL = Vitality	MHI = Mental Health
PAIN = Bodily Pain	SOC = Social Functioning	

The leflunomide treatment group showed significant improvements in 5 of the 8 SF-36 subscale scores compared to placebo and in two of the subscale scores compared to methotrexate. Statistically significant improvement compared to placebo was evident in physical functioning, pain, general health, vitality, and social role domains. Improvement in the physical functioning domain is consistent with the HAQ DI and PET results. Of interest, leflunomide-treated patients reported statistically significant improvements in bodily pain and vitality compared to methotrexate-treated patients. It has been noted that a change of 10 points in SF-36 is clinically meaningful [27]. Leflunomide subjects had improvements of over 10 points in 4 of the subscales, and almost 10 on 3 more (including social function and role emotional in addition to the physical domain scales).

1.3.2.4 Summary and conclusions for 6 and 12 month data (ITT)

In summary, statistically significant improvements in physical function (HAQ DI) and health-related quality of life (SF-36) were evident in leflunomide-treated patients; the improvements were statistically significantly superior to placebo and clinically meaningful. In the trials US301 and MN301, leflunomide results for the HAQ DI and several of its subscales were also statistically superior to those of the two active control medications (methotrexate and sulfasalazine, respectively). For US301, results of the PET showed leflunomide to be superior to placebo and methotrexate on activities important to the patient.

The active-controlled trial MN302 showed methotrexate to be statistically superior to leflunomide with regard to the HAQ DI. However, the difference between the two treatment groups was small (0.10) and may not reflect a clinically meaningful difference (0.22 according to [27]). Furthermore, the magnitude of change in the mean HAQ DI for the leflunomide group in MN302 was similar to that in US301 and MN301.

The results of the SF-36 PCS were in agreement with the results of the HAQ DI and PET in US301.

Leflunomide-treated patients clearly improved in the performance of essential activities of daily living as well as moderate and vigorous activities that are discretionary in nature, such as walking a block or mile, or climbing stairs.

1.3.3 Results for Year-2 Cohort

The year-2 cohort includes all patients initially enrolled into US301, MN301, and MN302 and who entered a 2nd year of treatment.

1.3.3.1 Demographic and baseline results for year-2 cohort

Study US301

The demographic and baseline characteristics of the year-2 cohort in US301 were similar to those of the ITT cohort (see table below). The disease duration in the small placebo group was somewhat longer in the year-2 cohort than in the ITT cohort (10.0 vs. 6.7 years).

Key demographic and baseline data for the ITT and year-2 cohorts of study US301

	ITT cohort			Year-2 cohort		
Characteristic	**Leflunomide (N=190)**	**Placebo (N=128)**	**Methotrexate (N=190)**	**Leflunomide (N=98)**	**Placebo (N=36)**	**Methotrexate (N=101)**
Female (%)	72.6	71.9	74.2	69.4	69.4	68.3
Mean age (years)	54.0	54.7	53.3	55.2	54.2	53.3
Age ≥ 65 years (%)	22.6	17.2	18.9	24.5	16.7	21.8
Race (%)						
Caucasian	88.9	89.1	89.5	92.9	88.9	92.1
Black	5.3	6.3	4.7	2.0	8.3	5.0
Asian	2.1	0	0.5	2.0	0	1.0
Other	3.7	4.7	5.3	3.1	2.8	2.0
Mean RA duration (years)	6.9	6.7	6.5	5.9	10.0	6.7
Mean number of prior DMARD treatments	0.8	0.9	0.9	0.8	0.9	0.9
Mean tender joint count	15.5	16.3	15.8	13.4	16.4	14.3
Mean swollen joint count	13.6	14.5	12.9	13.3	14.2	13.0

Study MN305

The demographic and baseline characteristics of the year-2 cohort in MN305 (i.e., patients entering Year 2 of blinded treatment) were similar to those of the ITT cohort (see table below).

Key demographic and baseline data, ITT and year-2 cohorts of study MN301/303/305

	ITT cohort		Year-2 cohort	
Characteristic	**Leflunomide (N=133)**	**Sulfasalazine (N=133)**	**Leflunomide (N=60)**	**Sulfasalazine (N=60)**
Female (%)	75.9	69.2	81.7	68.3
Mean age (years)	58.3	58.9	57.8	58.8
Age ≥65 years (%)	32.3	38.3	33.3	41.7
Race (%)				
White	86.5	93.2	86.7	90.0
Black	6.8	3.8	8.3	5.0
Other	6.8	3.0	5.0	5.0
Mean RA duration (years)	7.6	7.4	6.7	6.4
Mean number of prior DMARD treatments	1.2	1.0	1.0	0.7
Mean tender joint count	18.8	16.7	18.4	15.7
Mean swollen joint count	16.2	15.3	16.7	15.2

Study MN304

The demographic and baseline characteristics of the year-2 cohort (i.e., patients entering Year 2 of blinded treatment) in MN304 were similar to those of the ITT cohort of study MN302/304 (see table below).

Key demographic and baseline data, ITT and year-2 cohorts of study MN302/304

	ITT cohort		Year-2 cohort	
Characteristic	**Leflunomide (N=501)**	**Methotrexate (N=498)**	**Leflunomide (N=292)**	**Methotrexate (N=320)**
Female (%)	70.7	71.3	71.2	71.3
Mean age (years)	58.3	57.8	57.7	57.0
Age ≥65 years (%)	30.7	30.1	25.7	27.2
Race (%)				
White	98.8	98.6	99.3	98.8
Black	0.6	0.6	0.7	0.6
Other	0.6	0.8	0.0	0.6
Mean RA duration (years)	3.7	3.8	3.5	3.8
Mean number of prior DMARD treatments	1.1	1.1	1.0	1.1
Mean tender joint count	17.2	17.7	16.9	17.2
Mean swollen joint count	15.8	16.5	16.0	16.1

1.3.3.2 HAQ DI results for year-2 cohort (US301, MN305, and MN304)

1.3.3.2.1 HAQ DI and PET results for year-2 cohort in US301

Changes in HAQ DI and PET from baseline to month 24 (LOCF) for year-2 cohort in US301

The following table summarizes mean changes in HAQ DI and PET from baseline to month 24 in US301.

Study US301: year-2 cohort
Mean changes from baseline to month 24 (LOCF) in HAQ DI and PET weighted top 5 score

Parameter	Leflunomide	Methotrexate
HAQ DI		
No. of patients	97	101
Baseline mean	1.2	1.2
Mean change	–0.60*	–0.37
Mean % change	50%	31%
PET weighted top 5		
Baseline mean	19.9	18.4
Mean change	–9.12**	–4.34
Mean % change	46%	24%

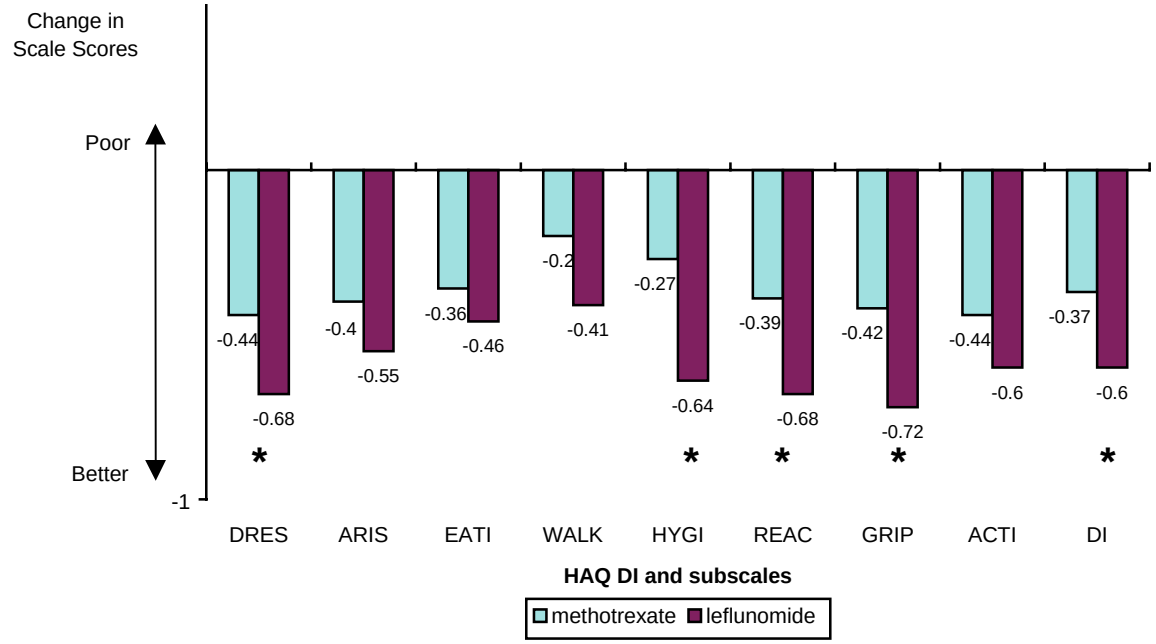
* Statistically significant difference between leflunomide and methotrexate, p=0.0050
** Statistically significant difference between leflunomide and methotrexate, p=0.0098

The mean change in HAQ DI showed statistically significant superiority of leflunomide compared to methotrexate at month 24. Improvement from baseline exceeded the 0.22-point threshold for clinical significance with both treatments, and the percentage change recommended by OMERACT [6], i.e. 36%, with leflunomide.

In the PET weighted score for the 5 activities most important to the patient, the leflunomide treatment group showed statistically significantly greater improvement compared to the methotrexate group. Therefore, improvements that were “customized” for each patient showed a statistically significantly greater improvement in the leflunomide group.

Changes between baseline and month 24 for the HAQ DI and its subscales in US301 are shown in the figure below.

Study US301: year-2 cohort
Mean changes from baseline to month 24 (LOCF) in HAQ DI and its subscales



* Leflunomide is significantly better at 0.05 level of significance.
Note: Number of patients varies between scales.

Scale abbreviations:

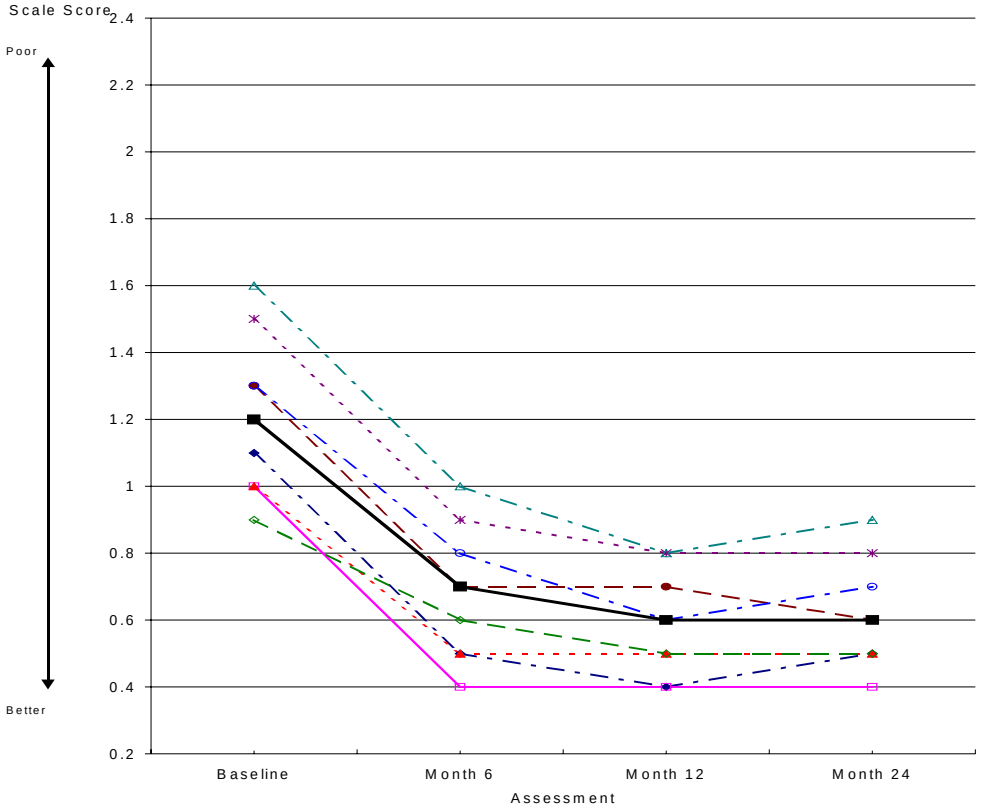
DRES	= Dressing	WALK	= Walking	GRIP	= Grip
ARIS	= Arising	HYGI	= Hygiene	ACTI	= Activities
EATI	= Eating	REAC	= Reach	DI	= Disability Index

At 24 months, the leflunomide treatment group showed statistically significant improvements when compared to methotrexate in HAQ DI and 4 of the 8 subscales, namely dressing, hygiene, reach, and grip. Furthermore, overall disability (as evidenced by the HAQ DI) showed a statistically significant improvement over methotrexate.

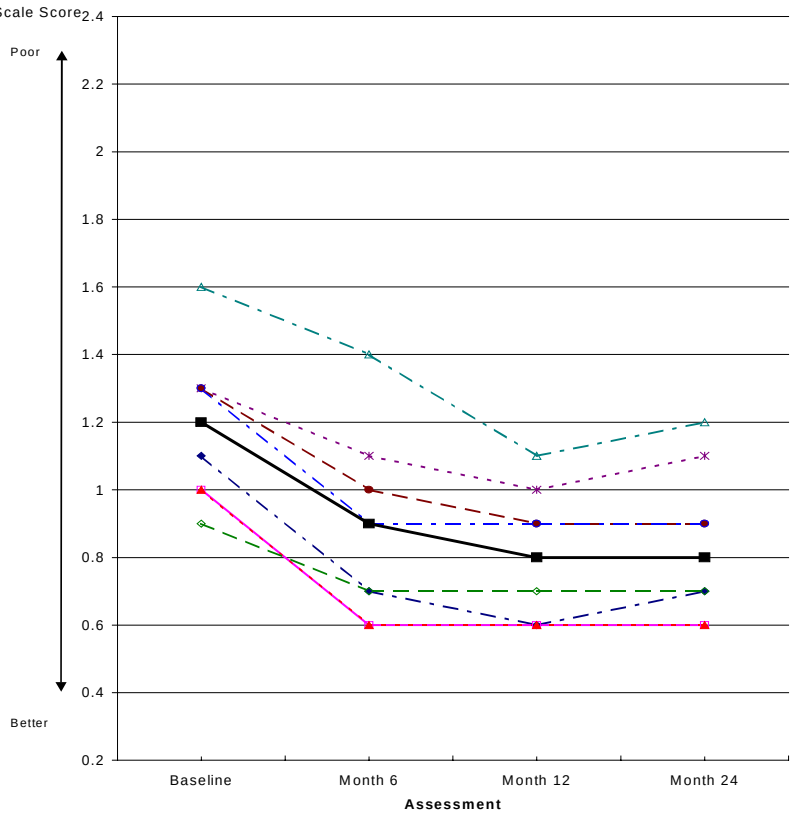
Maintenance of improvements in HAQ DI for year-2 cohort in US301 (LOCF)

The following diagrams show the scores for the HAQ DI and its subscales over time in leflunomide- and methotrexate-treated patients in US301.

Study US301: year-2 cohort
Mean HAQ DI and subscales over time in leflunomide-treated patients



Study US301: year-2 cohort
Mean HAQ DI and subscales over time in methotrexate-treated patients



Scale abbreviations:

DRES = Dressing
ARIS = Arising
EATI = Eating
WALK = Walking
HYGI = Hygiene
REAC = Reach
GRIP = Grip
ACTI = Activities
DI = Disability Index

The HAQ DI demonstrated that treatment effects for leflunomide and methotrexate were maintained between months 12 and 24. The 95% confidence intervals (CI) for the difference between month 12 and month 24 within treatment groups include zero and do not exceed the threshold for clinical relevance (0.22) as shown in the table below.

Summary of ANCOVA for HAQ DI in year-2 cohort of US301 (LOCF)

Treatment group	Mean difference (month 24–month 12)	95% CI	p-value
Leflunomide	0.014	(–0.047, 0.075)	0.6569
Methotrexate	0.013	(–0.067, 0.092)	0.7503

Therefore, the use of leflunomide or methotrexate for up to 24 months maintained physical function improvements that were demonstrated at 12 months and already evident after 6 months of treatment.

1.3.3.2.2 HAQ DI results for year-2 cohort in MN305

Changes in HAQ DI from baseline to month 24 (LOCF) for year-2 cohort in MN305

The table below summarizes mean changes from baseline to month 24 for the HAQ DI in MN305.

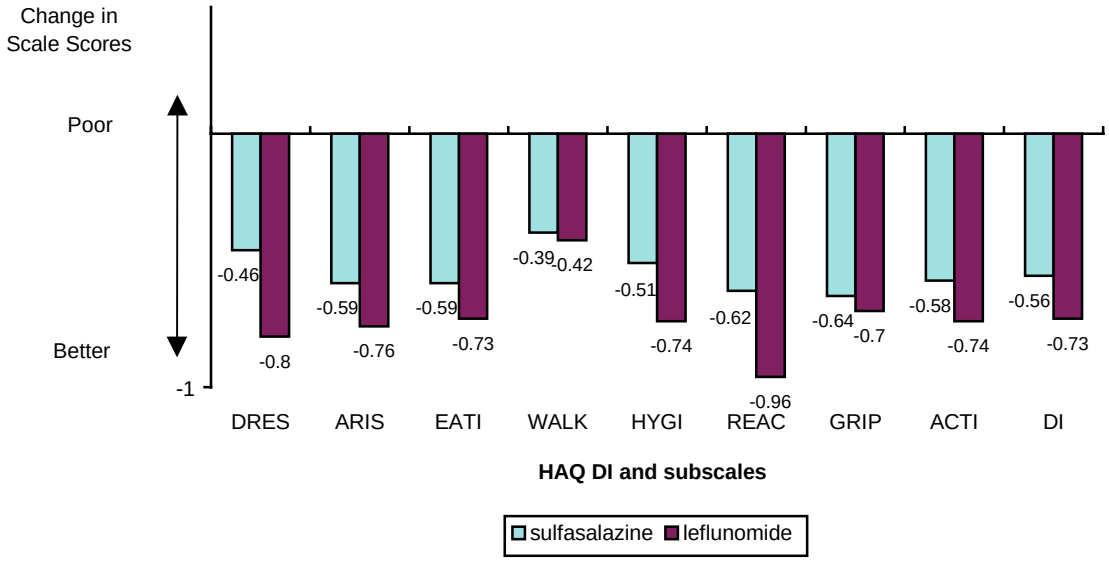
Study MN305: year-2 cohort
Mean change from baseline to month 24 in HAQ DI

Parameter	Leflunomide	Sulfasalazine
No. of patients	51	46
Baseline mean	1.6	1.5
Mean change	–0.73	–0.56
Mean % change	46%	37%

The changes in the HAQ DI for the leflunomide and sulfasalazine treatment groups clearly exceeded the clinically meaningful threshold of –0.22 and those recommended by OMERACT (36%). The mean change in HAQ DI showed no statistically significant difference between the two treatment groups.

Changes between baseline and month 24 for the HAQ DI and its subscales in MN305 are shown in the figure below.

Study MN305: year-2 cohort
Mean changes from baseline to month 24 (LOCF) in HAQ DI and its subscales



Note: Number of patients varies between scales. There were no significant differences between leflunomide and sulfasalazine.

Scale abbreviations:

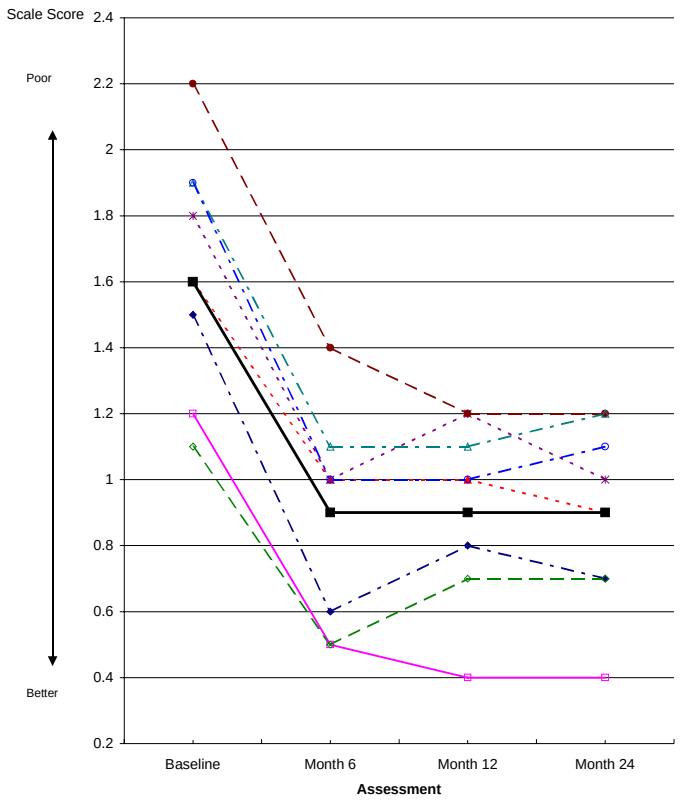
DRES	= Dressing	WALK	= Walking	GRIP	= Grip
ARIS	= Arising	HYGI	= Hygiene	ACTI	= Activities
EATI	= Eating	REAC	= Reach	DI	= Disability Index

Patients in the leflunomide treatment group demonstrated a greater response than the sulfasalazine group on all subscales of the HAQ DI. Changes for all subscales except walking exceeded –0.7, indicating clinically meaningful changes. These data support the same trend seen with the year-1 data.

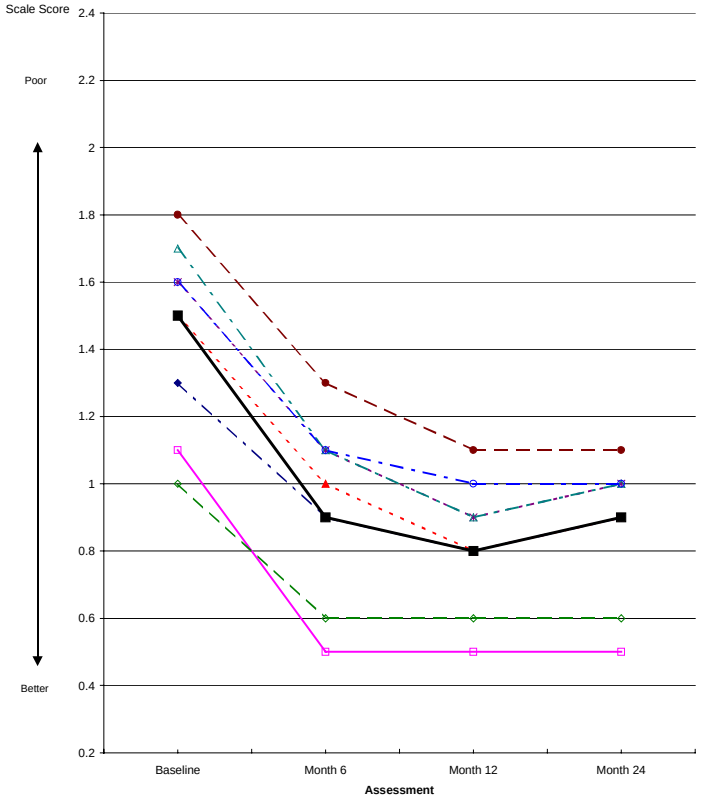
Maintenance of improvements in HAQ DI for year-2 cohort in MN305 (LOCF)

The following diagrams show the scores for the HAQ DI and its subscales over time in leflunomide- and sulfasalazine-treated patients in MN305.

Study MN305: year-2 cohort
Mean HAQ DI and subscales over time in leflunomide-treated patients



Study MN305: year-2 cohort
Mean HAQ DI and subscales over time in sulfasalazine-treated patients



Scale abbreviations:

DRES = Dressing
ARIS = Arising
EATI = Eating
WALK = Walking
HYGI = Hygiene
REAC = Reach
GRIP = Grip
ACTI = Activities
DI = Disability Index

The HAQ DI demonstrated that treatment effects for leflunomide and sulfasalazine were maintained between months 12 and 24. The 95% CIs for the difference between month 12 and month 24 within treatment groups include zero and do not exceed the threshold for clinical relevance (0.22) as shown in the table below.

Summary of ANCOVA for HAQ DI in year-2 cohort of MN305 (LOCF)

Treatment group	Mean difference (month 24–month 12)	95% CI	p-value
Leflunomide	–0.028	(–0.105, 0.050)	0.4775
Sulfasalazine	0.060	(–0.035, 0.154)	0.2090

Therefore, the use of leflunomide or sulfasalazine for up to 24 months maintained physical function improvements that were demonstrated at 12 months and already evident after 6 months of treatment.

1.3.3.2.3 HAQ DI results for year-2 cohort in MN304

Changes in HAQ DI from baseline to month 24 (LOCF) for year-2 cohort in MN304

The table below summarizes mean changes from baseline to month 24 for the HAQ DI in MN304.

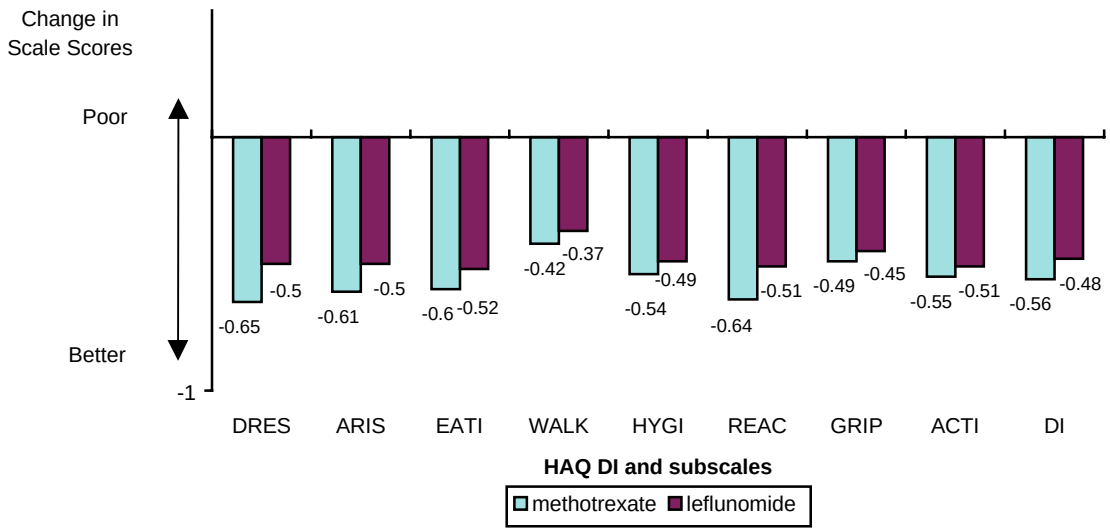
Study MN304: year-2 cohort
Mean change from baseline to month 24 (LOCF) in HAQ DI

Parameter	Leflunomide	Methotrexate
No. of patients	248	273
Baseline mean	1.5	1.5
Mean change	–0.48	–0.56
Mean % change	32%	37%

The changes in the HAQ DI for the leflunomide and methotrexate treatment groups clearly exceeded the clinically meaningful threshold of –0.22. T he mean change in HAQ DI showed no statistically significant difference between the two treatment groups.

Changes between baseline and month 24 for the HAQ DI and its subscales in MN304 are shown in the figure below.

Study MN304: year-2 cohort
Mean changes from baseline to month 24 (LOCF) in HAQ DI and its subscales



Note: Number of patients varies between scales. There were no significant differences between leflunomide and methotrexate.

Scale abbreviations:

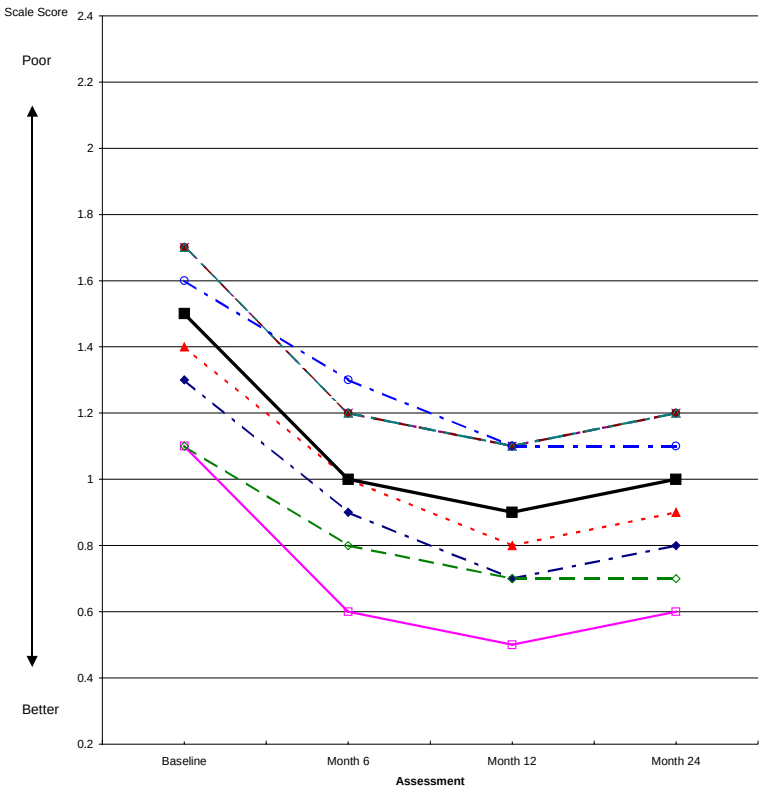
DRES = Dressing	WALK = Walking	GRIP = Grip
ARIS = Arising	HYGI = Hygiene	ACTI = Activities
EATI = Eating	REAC = Reach	DI = Disability Index

The changes for each scale in the leflunomide treatment group approached or exceeded –0.4, representing clinically important differences.

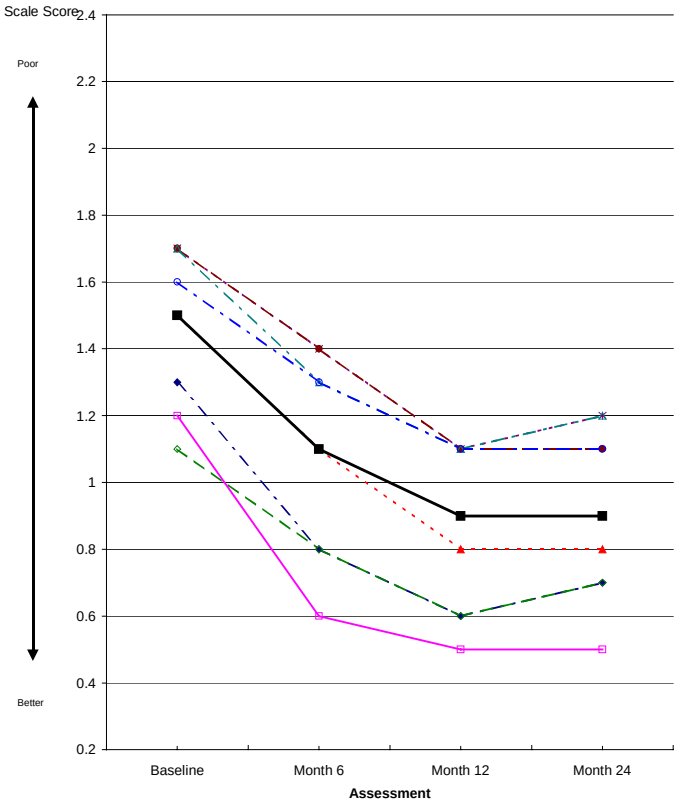
Maintenance of improvements in HAQ DI for year-2 cohort in MN304 (LOCF)

The following diagrams show the scores for the HAQ DI and its subscales over time in leflunomide- and methotrexate-treated patients in MN304.

Study MN304: year-2 cohort
Mean HAQ DI and subscales over time in leflunomide-treated patients



Study MN304: year-2 cohort
Mean HAQ DI and subscales over time in methotrexate-treated patients



Scale abbreviations:

DRES = Dressing
ARIS = Arising
EATI = Eating
WALK = Walking
HYGI = Hygiene
REAC = Reach
GRIP = Grip
ACTI = Activities
DI = Disability Index

In the leflunomide group, physical function improvements demonstrated at 12 months were maintained at 24 months in 4 of the 8 subscales, namely arising, eating, walking, and activities. Improved functioning for all scales was evident following 6 months of treatment. This clinically meaningful change was also evident following 12 and 24 months of treatment. The HAQ DI and the remaining 4 subscale scores increased; however, the increases were small and do not represent clinically meaningful changes. This is supported by the fact that the 95% CI for the mean change in HAQ DI (0.029, 0.140) does not include zero (statistically significant difference); however, it excludes 0.22, the change needed for clinically important differences.

In the methotrexate group, physical function improvements demonstrated at 12 months were maintained at 24 months in the HAQ DI and all subscales except walking, which increased.

Summary of ANCOVA for HAQ DI in year-2 cohort of MN304 (LOCF)

Treatment group	Mean difference (month 24–month 12)	95% CI	p-value
Leflunomide	0.084	(0.029, 0.140)	0.0032
Methotrexate	0.043	(–0.008, 0.094)	0.0999

1.3.3.3 SF-36, MOS Current Health, and WPQ for year-2 cohort (US301 only)

Changes in SF-36, MOS Current Health, and WPQ from baseline to month 24 (LOCF) for year-2 cohort in US301

Changes in the SF-36 PCS, SF-36 MCS, MOS Current Health, and WPQ in US 301 are summarized in the table below.

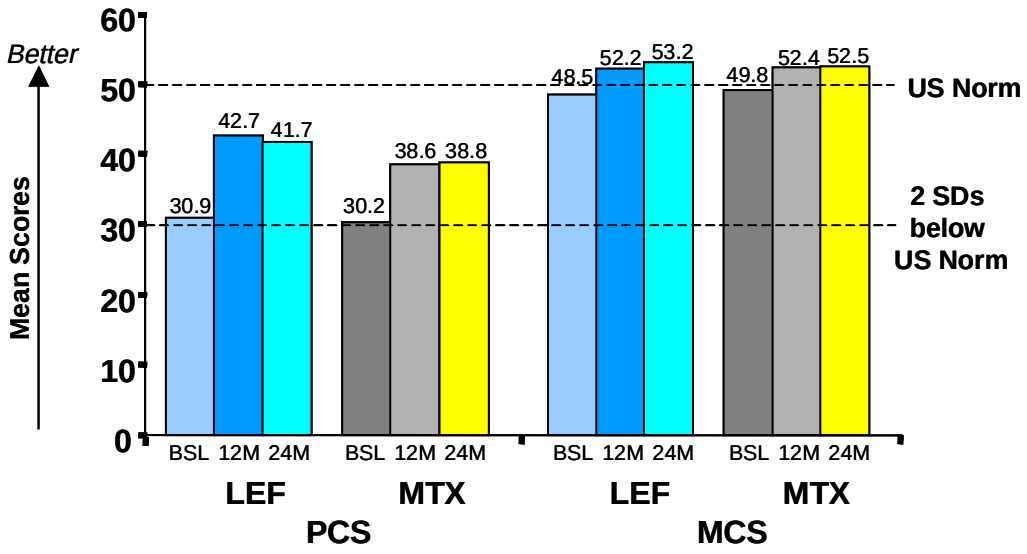
Study US301: year-2 cohort
Mean changes from baseline to month 24 (LOCF) for SF-36 PCS, SF-36 MCS, MOS Current Health, and WPQ

Parameter	Leflunomide	Methotrexate
SF-36 PCS		
No. of patients	93	97
Baseline mean	30.9	30.2
Mean change	10.8	8.4
Mean % change	35%	28%
SF-36 MCS		
No. of patients	93	97
Baseline mean	48.5	49.8
Mean change	4.7	2.7
Mean % change	10%	5%
MOS Current Health		
No. of patients	91	97
Baseline mean	49.2	43.8

Parameter	Leflunomide	Methotrexate
Mean change	16.8	17.3
Mean % change	34%	39%
Work Productivity Questionnaire		
No. of patients	74	77
Baseline mean	54.5	54.7
Mean change	14.2	11.5
Mean % change	26%	21%

The changes from baseline to month 24 in the SF-36 PCS, SF-36 MCS, MOS Current Health score, and WPQ showed no statistically significant differences between the leflunomide and methotrexate treatment groups. Improvement in PCS and MCS Scores for year-2 cohort patients receiving leflunomide and methotrexate is shown in the figure below.

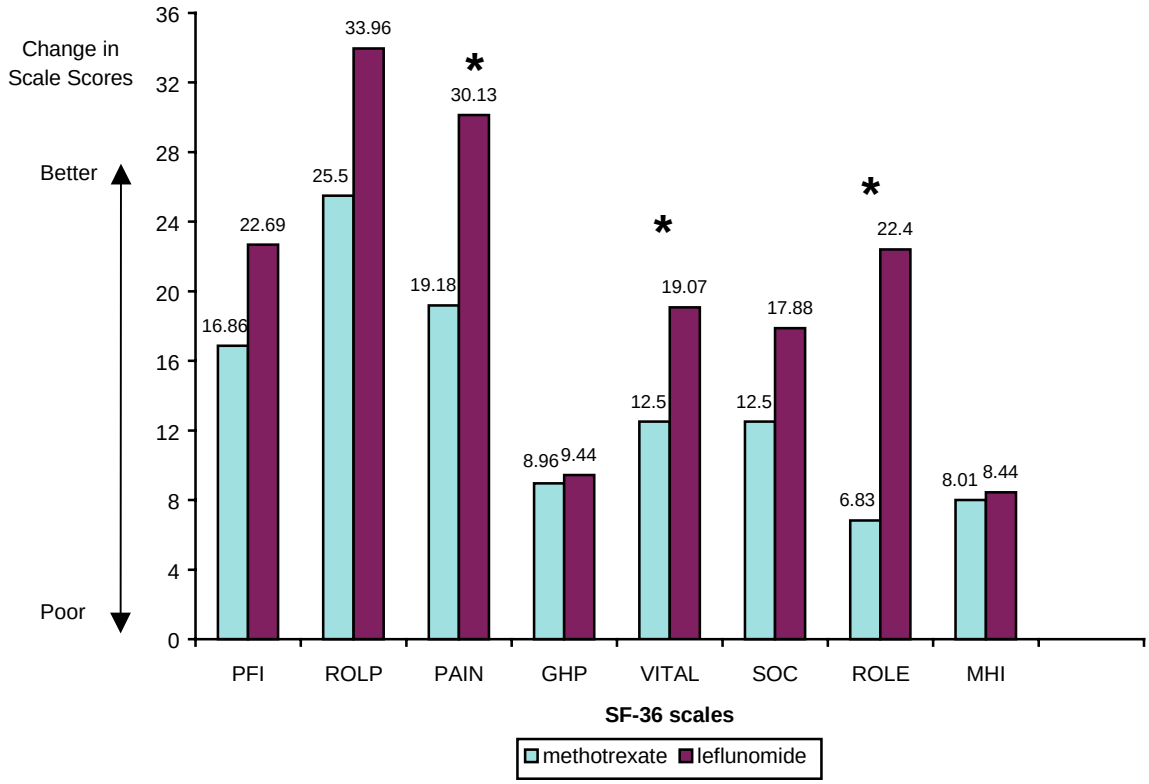
US301: Improvement in PCS and MCS Scores
Leflunomide and Methotrexate: Year-2 Cohort



PCS= Physical Component Summary Score
MCS=Mental Component Summary Score

Mean changes in the subscale scores of the SF-36 at month 24 are shown in the figure below.

Study US301: year-2 cohort
Mean changes from baseline to month 24 (LOCF) in SF-36 scales



* Leflunomide is significantly better at 0.05 level of significance.
Note: Number of patients varies between scales.

Scale abbreviations:

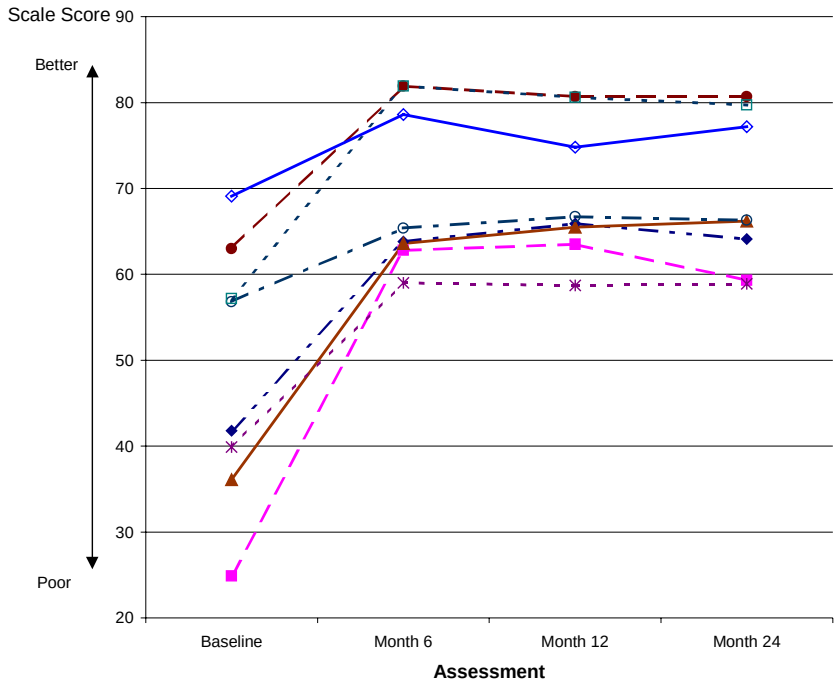
PFI = Physical Functioning	GHP = General Health Perception	ROLE = Role Emotional
ROLP = Role Physical	VITAL = Vitality	MHI = Mental Health
PAIN = Bodily Pain	SOC = Social Functioning	

The leflunomide treatment group showed significantly greater improvements compared to methotrexate in the three following SF-36 subscales: bodily pain, vitality, and role emotional.

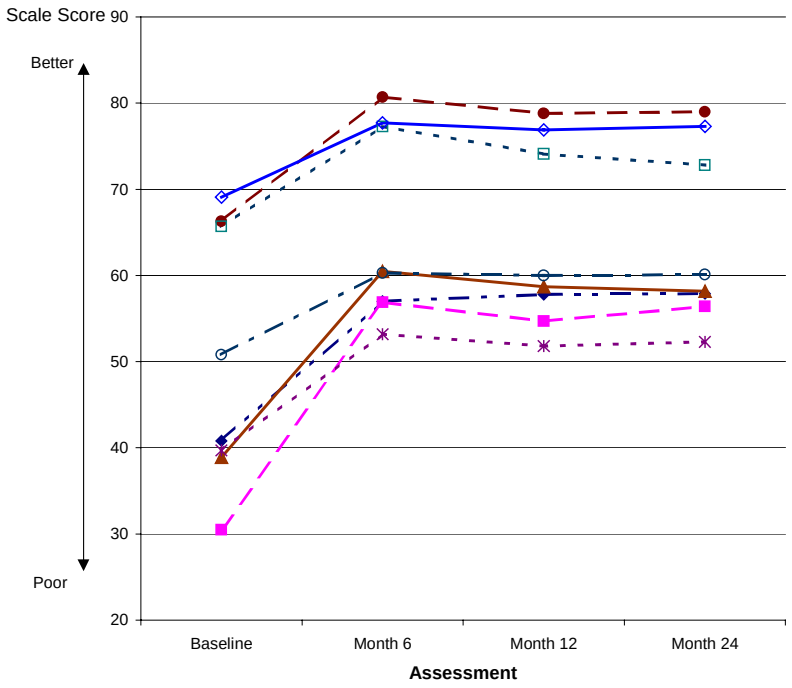
Maintenance of improvements in SF-36 for year-2 cohort in US301 (LOCF)

The following diagrams show the SF-36 subscale scores over time for leflunomide- and methotrexate-treated patients in US301.

Study US301: year-2 cohort
Mean SF-36 scales over time in leflunomide-treated patients



Study US301: year-2 cohort
Mean SF-36 scales over time in methotrexate-treated patients



Scale abbreviations:

PFI = Physical Functioning
GHP = General Health Perception
ROLE = Role Emotional
ROLP = Role Physical
VITAL = Vitality
MHI = Mental Health
PAIN = Bodily Pain
SOC = Social Functioning

The improvement in health-related quality of life, as measured by the SF-36 scales, demonstrated that treatment effects for leflunomide and methotrexate were maintained between months 12 and 24. The 95% CIs for the summary scales include zero and do not exceed the threshold for clinical relevance (10 points [20]) as shown in the table below.

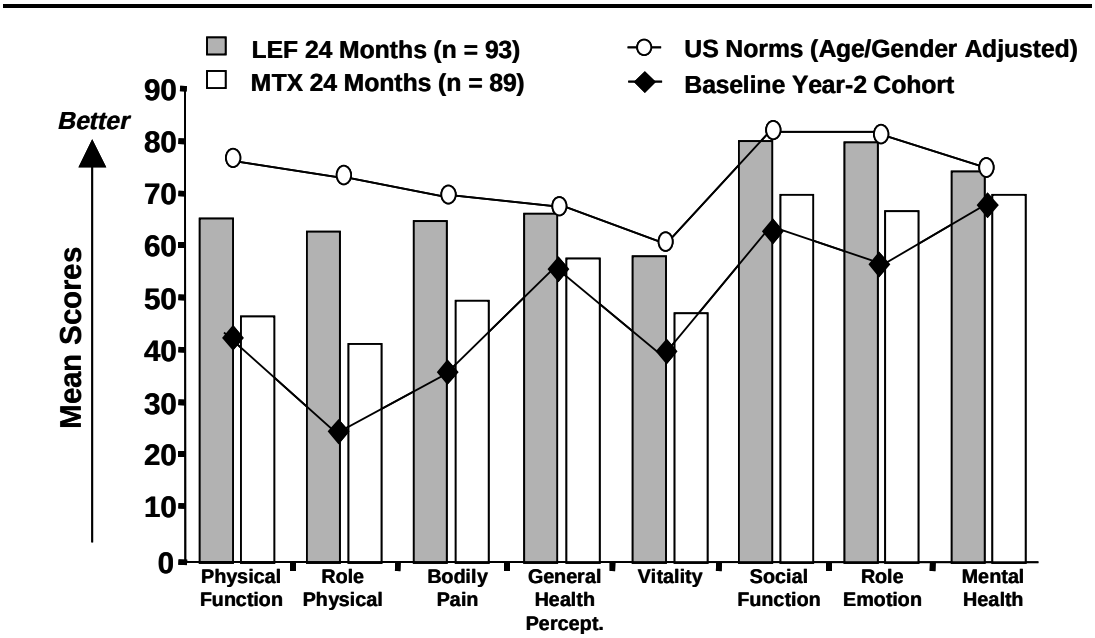
Summary of ANCOVA for SF-36 summary scores in year-2 cohort of US301 (LOCF)

Treatment group	Mean difference (month 24–month 12)	95% CI	p-value
SF-36 PCS			
Leflunomide	0.971	(–2.392, 0.450)	0.1781
Methotrexate	0.244	(–1.185, 1.673)	0.7353
SF-36 MCS			
Leflunomide	0.726	(–0.863, 2.316)	0.3666
Methotrexate	0.106	(–1.212, 1.425)	0.8731

Therefore, the use of leflunomide or methotrexate for up to 24 months maintained improvements in health-related quality of life (as measured by the SF-36 scales) that were demonstrated at 12 months and already evident after 6 months of treatment.

Compared with age and gender-adjusted US norms, improvements in the year-2 cohort approached normative values in the leflunomide treatment group, as shown in the following figure.

US301 Year-2 Cohorts: Mean Improvement in SF-36 Leflunomide and Methotrexate



1.3.3.4 Summary and conclusions for year-2 cohort

Treatment with leflunomide demonstrated statistically significant and clinically meaningful changes at the end of 1 year of treatment for the HAQ DI in all three studies as well as for PET, SF-36, MOS Current Health, and WPQ in US301. The improvements in HAQ DI seen in patients treated with leflunomide over 24 months represent clinically meaningful improvements that are twice the recommended threshold of –0.22.

The leflunomide treatment group statistically significantly demonstrated the maintenance of improved physical function (US301 and MN305) and health-related quality of life (US301) over 2 years and in several function scales of MN304. There were no clinically meaningful changes between month 12 and month 24 within any of the leflunomide groups.

These data support the efficacy of leflunomide with regard to physical function and health-related quality of life.

1.3.4 2-Year Results (ITT Population)

1.3.4.1 LOCF analysis

The changes in HAQ DI and SF-36 summary scores from baseline to month 24 within each treatment group are summarized in the following table.

Mean changes from baseline to month 24 for HAQ DI and SF-36 summary scores in Phase III studies: ITT population

Parameter	Leflunomide	Placebo	Sulfasalazine	Methotrexate
HAQ DI (US301)				
No. of patients	179	121	–	179
Baseline mean	1.3	1.4	–	1.3
Mean change	–0.427*, **	–0.062	–	–0.256
HAQ DI (MN305)				
No. of patients	114	81	111	–
Baseline mean	1.649	1.591	1.496	–
Mean change	–0.577*, **	–0.078	–0.374	–
HAQ DI (MN304)				
No. of patients	462	–	–	457
Baseline mean	1.503	–	–	1.522
Mean change	–0.411	–	–	–0.522**
SF-36 MCS (US301)				
No. of patients	168	116	–	171
Baseline mean	46.7	47.7		47.9
Mean change	1.971	1.213	–	1.423
SF-36 PCS (US301)				
No. of patients	168	116	–	171
Baseline mean	30.1	29.2	–	29.4
Mean change	7.069*, **	1.825	–	4.831
PET (US301)				
No. of patients	179	122	-	180
Baseline mean	20.8	22.7	-	20.6
Mean change	-6.3*,**	-1.4	-	-3.4

* Indicates statistically significant differences between leflunomide and placebo p<0.001.
** Indicates statistically significant differences between leflunomide and methotrexate or sulfasalazine p<0.05.

The following table summarizes the ANCOVA analysis of differences between treatment groups in change from baseline to month 24 for the ITT population using LOCF.

Summary of ANCOVA for change from baseline to month 24 in ITT populations of Phase III studies (LOCF)

Variable/treatment comparison	Mean adjusted difference	95% CI	p-value
HAQ DI (US301)*			
LEF – PBO	–0.397	(–0.539,–0.255)	<0.0001
LEF – MTX	–0.236	(–0.381,–0.091)	0.0015
MTX – PBO	–0.164	(–0.305,–0.022)	0.0233
HAQ DI (MN305)*			
LEF – PBO	–0.497	(–0.651, –0.344)	<0.0001
LEF – SSZ	–0.175	(–0.330, –0.019)	0.0279
HAQ DI (MN304)*			
LEF – MTX	0.106	(0.028, 0.183)	0.0077
SF-36 MCS (US301)**			
LEF – PBO	0.554	(–2.328, 3.435)	0.7056
LEF – MTX	0.427	(–2.196, 3.050)	0.7491
MTX – PBO	0.161	(–2.429, 2.751)	0.9027
SF-36 PCS (US301)**			
LEF – PBO	6.581	(3.725, 9.437)	<0.0001
LEF – MTX	3.955	(0.963, 6.946)	0.0097
MTX – PBO	2.729	(–0.122, 5.580)	0.0606

LEF = leflunomide, MTX = methotrexate, PBO = placebo, SSZ = sulfasalazine

***** For HAQ DI, a negative adjusted mean difference represents an advantage for the first of the two treatment groups listed in the first column.

****** For the SF-36 summary scores, a positive adjusted mean difference represents an advantage for the first of the two treatment groups listed in the first column.

Leflunomide treatment demonstrated statistically significantly greater improvements in HAQ DI than placebo in studies US301 and MN301/303/305. In these studies statistically significantly greater improvement also occurred in the leflunomide groups compared to methotrexate or sulfasalazine. Conversely, results in MN304 showed a statistically significant greater improvement in the methotrexate group compared to leflunomide, although notable improvements occurred in both active treatment groups.

It has been noted that a change of –0.22 points in the HAQ DI (i.e., an improvement of 0.22) reflects a clinically meaningful change [27]. In the leflunomide treatment group, all measures met or exceeded this level: the mean HAQ DI improved by 0.44 to 0.67, reflecting a clinically meaningful change in all studies.

In study US301, leflunomide was statistically significantly superior to placebo and methotrexate in the SF-36 PCS.

In summary, the analysis of the ITT population at month 24 (LOCF) demonstrated statistically significant and clinically meaningful improvements in physical function and health-related quality of life for leflunomide compared to placebo.

1.3.4.2 Sensitivity analysis

1.3.4.2.1 Methodology of sensitivity analysis

Missing data due to withdrawals is one of the major issues in the analysis of controlled clinical trials and is especially important in long-term, placebo-controlled trials. A sensitivity analysis was, therefore, performed to demonstrate the robustness of the treatment effects shown in the year-2 and ITT (LOCF) analyses described in *Sections 1.3.2,1.3.3,* and *1.3.4.1*.

Furthermore, sensitivity analysis based on the subset of patients who continued in the long-term therapy protocols evaluates the plausibility of clinical outcomes that would be required to invalidate statistically significant results.

The sensitivity analysis employed repeated random sampling from the non-missing cases to replace missing data from the same treatment group in the ITT population. The three following non-missing cohorts were used for sampling purposes:

1. Completers with month 24 data: Patients completing 24 months of treatment and with HAQ DI and SF-36 PCS data at week 104.
2. Completers with LOCF: Patients completing 24 months of treatment using LOCF for those without HAQ DI or SF-36 PCS data at month 24.
3. Patients with data at exit visit: Patients with HAQ DI or SF-36 PCS assessment at their study exit visit. These patients could have completed 24 months of treatment, dropped out of the study earlier, or switched to alternate therapy. Data from this subset do not include results from the alternate therapy phase.

The sensitivity analysis presented in this *Briefing Document* evaluated the boundary treatment effect of the cohort without HAQ DI and SF-36 PCS data at month 24 such that statistical significance would disappear in the ITT population. A judgment may then be made as to whether such a boundary effect in the missing data cohort is plausible and realistic.

1.3.4.2.2 Results of sensitivity analysis

The following table summarizes the results of the sensitivity analysis for the HAQ DI in US301.

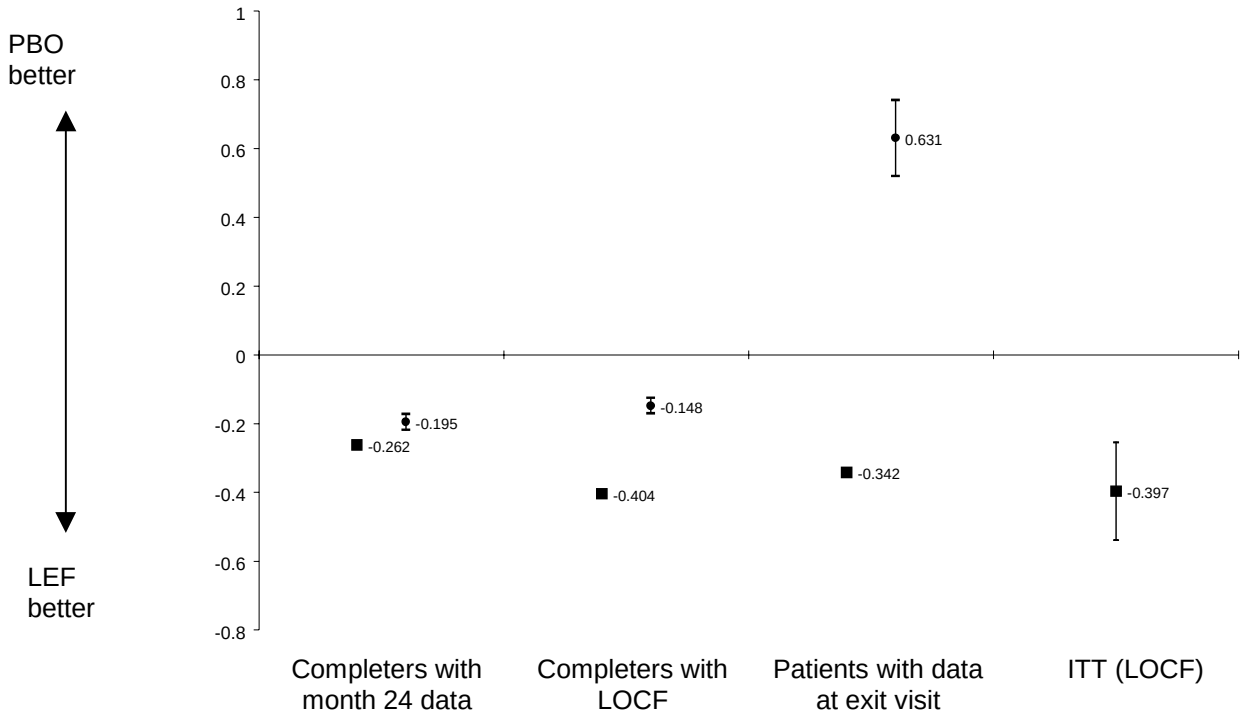
Sensitivity analysis of HAQ DI in US301

	Analysis approach (type of non-missing cohort)		
Parameter	**Completers with month 24 data**	**Completers with LOCF**	**Patients with data at exit visit**
Total patients/patients with non-missing data	318/78	318/110	318/255
Mean change from baseline to month 24			
Leflunomide	–0.573	–0.656	–0.369
Placebo	–0.311	–0.252	–0.027
Treatment difference	–0.262	–0.404	–0.342
Sensitivity analysis *			
Average sampled mean treatment difference	–0.208	–0.322	–0.335
Average sampled p-value	0.0028	<0.0001	<0.0001
Average boundary values required in missing cohort to maintain p≤0.05 (95% CI)**	–0.195 (–0.218, –0.172)	–0.148 (–0.170, –0.125)	0.631 (0.520, 0.741)

* Missing data replaced with data from non-missing cohort.
** Mean treatment difference given as least squares (LS) mean.

The results of the sensitivity analysis for the HAQ DI in US301 are illustrated in the following diagram.

Comparison of effects in three non-missing cohorts and ITT (LOCF) for HAQ DI in US301



Note: The squares show the point estimates for the treatment differences between placebo and leflunomide in the non-missing cohorts. The circles show the average boundary values (with 95% CIs) of the missing cohort needed to maintain significance ($p<0.05$) for treatment differences.

The point estimates for the treatment differences in the non-missing cohorts all lie within the 95% CI for the difference between placebo and leflunomide in the ITT population.

The diagram shows the 95% CIs for the boundary values that would need to be achieved in each of the three missing cohorts to make the difference in the overall cohort statistically non-significant. Of the three cohorts, the cohort of patients with data at exit visit has the smallest proportion of missing data and the results for these missing data would have to contradict all other data (i.e., placebo would have to be better than leflunomide) to achieve an overall non-significant result.

The following table summarizes the results of the sensitivity analysis for the SF-36 PCS in US301.

Sensitivity analysis of SF-36 PCS in US301

	Analysis approach (type of non-missing cohort)		
Parameter	**Completers with month 24 data**	**Completers with LOCF**	**Patients with data at exit visit**
Total patients/patients with non-missing data	318/80	318/110	318/248
Mean change from baseline to month 24			
Leflunomide	10.607	11.670	6.582
Placebo	11.955	3.784	1.435
Treatment difference	–1.348	7.885	5.147
Sensitivity analysis *			
Average sampled mean treatment difference	3.730	6.411	4.967
Average sampled p-value	0.0163	0.0003	0.0008
Average boundary values required in missing cohort to maintain p≤0.05 (95% CI)**	1.774 (1.374, 2.173)	1.342 (1.179, 1.505)	–3.484 (–4.172, 2.797)

* Missing data replaced with data from nonmissing cohort.
** Mean treatment difference given as least squares (LS) mean.

Non-missing cohort – completers with month 24 data

When the non-missing cohort is defined as completers with month 24 data, the differences between the leflunomide and placebo groups in HAQ DI or SF-36 PCS were not statistically significant due to the small sample sizes. In the leflunomide treatment group, the number of patients in the missing cohort was more than double (132/58=2.3) that in the non-missing cohort. In the placebo group, the number of patients in the missing cohort was more than five times (108/20=5.4) that in the nonmissing cohort.

When sensitivity analysis was performed by replacing missing data with data from the non-missing cohort, the differences were statistically significant in all iterations. If the missing cohort were behaving similarly to the non-missing cohort, the difference between leflunomide and placebo would have been statistically significant in the overall population.

Non-missing cohort – completers using LOCF

When the non-missing cohort is defined as completers using LOCF, the differences between the two treatment groups in HAQ DI or SF-36 PCS were statistically significant (p=0.0240 and p=0.0319, respectively).

When sensitivity analysis was performed by replacing the missing data with data from the non-missing cohort, the differences were statistically significant in all iterations. If the missing cohort were behaving similarly to the non-missing cohort, the difference between leflunomide and placebo would have been statistically significant in the overall population.

Non-missing cohort – all patients with exit visit data

When the non-missing cohort is defined as all patients with HAQ DI or SF-36 PCS data at their exit visit, the differences between leflunomide and placebo were statistically significant (p=0.0001 for HAQ DI and p=0.0021 for SF-36 PCS).

When sensitivity analysis was performed by replacing the missing data with data from the non-missing cohort, the differences were statistically significant in all iterations. Furthermore, the leflunomide performance in the missing cohort would have to be very much worse than that of placebo in order to nullify the statistically significant difference in the overall population, which is not clinically plausible.

1.3.4.2.3 Conclusions from sensitivity analysis

The sensitivity analysis used three approaches to replace missing data and showed superiority of leflunomide over placebo for the HAQ DI and SF-36 PCS. It thus demonstrated the robustness of the year-2 cohort and ITT LOCF analyses.

1.3.5 Clinical Relevance

Minimum clinically important difference (MCID)

Mean or median improvements in a treatment group that are statistically significant compared with placebo frequently are not necessarily clinically meaningful or readily understood. Recent efforts designed to develop consensus regarding outcome measures in clinical trials have included discussion of “minimum clinically important differences” [MCID], e.g. degrees of improvement in various outcome measures that would be perceptible to patients, on an individual basis, and would be considered clinically meaningful to them. Improvements of 33 to 36% over baseline (or 18% greater than placebo) are thought to be clinically important [1,2]. Although these definitions are relevant only on an individual patient basis, when mean and median changes within a treatment group well exceed such a value it can be estimated that the majority of the group will have attained clinically important improvements.

In recent longitudinal and randomized controlled trials in RA, OA and SLE, as well as chronic cardiovascular and pulmonary conditions, changes in a variety of patient assessed outcome measures including global assessments of disease activity/severity as well as pain and physical function have been correlated with changes observed in domains of the SF-36 as well as PCS and MCS summary scores.

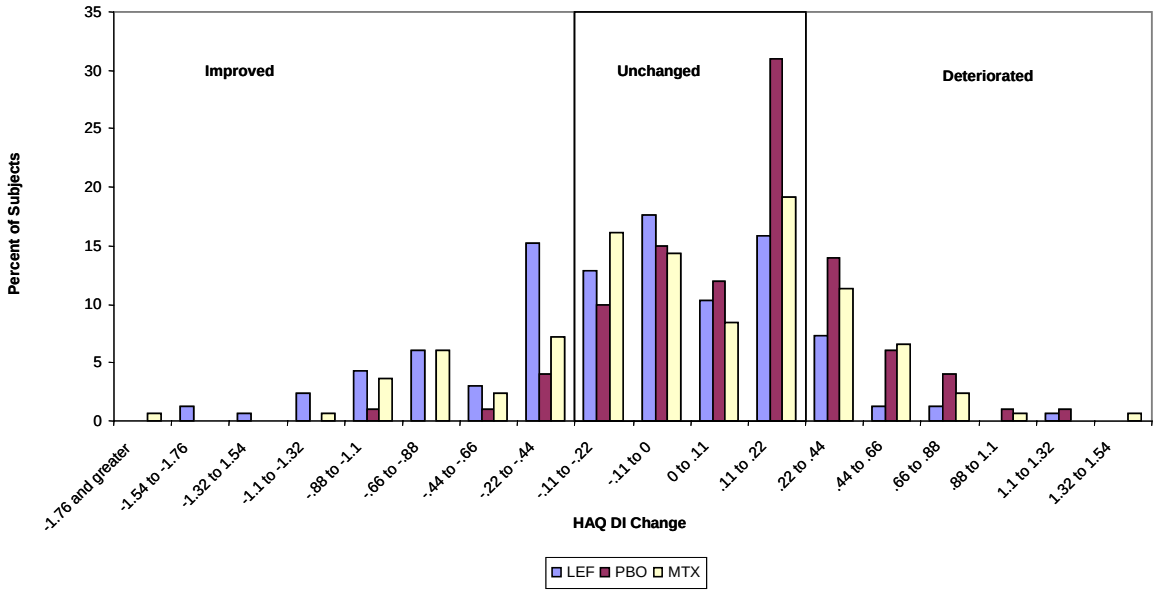
Wyrwich et al. compared the standard error of measurement [SEM] in SF-36 domains to MCID differences in the Chronic Heart Failure Questionnaire in one RCT, and the Chronic Respiratory Disease Questionnaire in another [30, 31]. In both studies, a value of one SEM in change scores for the SF-36 domains closely approximated MCIDs for the disease specific questionnaire components. The SEMs for SF-36 domain change scores ranged from 7.88 to 15.26 in the first comparison and 7.65 to 14.15 in the second.

Kosinski and Ware compared changes in Health Assessment Questionnaire disability index [HAQ DI] and SF-36 domains and summary scores with patient global assessments and pain in two RCTs comparing COX-2 selective agents to traditional NSAIDs in active RA [9]. Mean changes in SF-36 domain scores corresponding to one level of improvement in patient global assessment or pain ranged from 4.2 to 21.0, and 1.9 to 10.8; 4.4 and 3.0 for PCS and 4.7 and 2.2 for MCS summary scores. Using the same technique to evaluate the HAQ DI yielded good agreement [-0.24 to -0.22] with previously published values for MCID of -0.22 [11]. Kujawski, Thumboo, Ehrich, Stucki and others have suggested that changes of 5 to 10 points in domain and 2.5 to 5 points in PCS and MCS summary scores are associated with meaningful clinical improvements and can be considered to represent MCID in RA, SLE and OA [5,6,21,32,33,34,35,36,37,38].

Analyses of clinical trial data from a variety of therapeutic agents in RA have indicated that improvements of 0.19 to 0.43 in HAQ DI scores correlate with ACR response rates and represent clinically meaningful changes, or MCID [7,9,14,27]. Similar statistical analyses have suggested that 5 points (or 33%) improvement in the PET top 5 scores are clinically important 7,27]. Although minimum clinically important differences in SF-36 have not yet been formally defined, Ware, Kosinski, Thumboo and others have suggested that changes of 5 to 10 points in domain and summary scores are associated with meaningful clinical improvements [8,9.10,16,20].

Analysis of the distribution of HAQ DI in US301 at 24 months revealed greater improvements (as defined by the percentage of patients whose score changes exceeded the MCID threshold) with leflunomide than with methotrexate or placebo, as shown in the figure below.

HAQ DI Distribution at 24 Months, Study US301



Mean changes in the HAQ DI in the leflunomide treatment groups over 24 months were consistent across the three Phase III trials. Improvements ranged from –0.48 to –0.73, more than twice the published values considered to represent MCID (0.19 to 0.22), corresponding to changes of 32% to 50% from baseline. They indicate that the majority of patients achieved clinically meaningful improvements, as summarized in the following table.

Percentages of leflunomide patients achieving MCID in HAQ DI

Cohort/study	6 months	12 months	24 months
ITT population			
US301	71%	71%	61%
MN305	68%	70%	70%
MN304	57%	62%	60%
Year-2 cohort			
US301	74%	76%	71%
MN305	76%	78%	80%
MN304	64%	72%	67%

Improvements in upper and lower extremity function, as evidenced by HAQ DI and its individual subscales indicated that patients in the leflunomide treatment groups improved in:

- dressing and grooming without aid
- arising from a seated position without aid
- cutting food, eating, and drinking without aid
- walking without aid
- maintaining self-hygiene and using a toilet without aid
- reaching and picking up objects from above or from the floor without aids
- strengthened grip
- performing activities (e.g., errands, chores) without aid, and
- overall disability

Improvement in the PET top 5 scores in the leflunomide treatment group in US301 was –9.12 from a baseline of 19.9, representing a 46% change. This would be considered well above the estimated value for MCID. Approximately 40% of the patient population in this study had early disease (≤2 years duration) and/or were DMARD naïve. Data from the COBRA trial indicated that, although the HAQ remains the instrument of choice in clinical trials in RA due to its ease of use, the MACTAR (or PET) is particularly responsive to change in patients with early disease [24]. In US301, the PET top 5 score indicated that reported changes in physical function were meaningful to patients, improving performance of those physical activities they considered important and most wanted changed.

In US301, where a generic measure of health-related quality of life was utilized, the leflunomide treatment group demonstrated statistically significant improvements in SF-36 domains and the PCS score compared with placebo and methotrexate after 1 year of treatment. Ruta et al reported that 4 of the SF-36 domains (pain, vitality, social function and physical function) were most responsive to change in their cohort of patients with active RA [15]. Similarly, results from US301 demonstrated changes in the same 4 domains, as well as improvements in the remaining 4 domains. Mean changes over 24 months in all SF-36 domains approximated or exceeded 10 points, ranging from 8.4 in mental

health index and 9.4 in general health profile to 19.1 in vitality, 22.4 in role emotional, 30.1 in pain, and 34.0 in role physical. These represent several multiples of the proposed MCID values for SF-36. Improvements maintained over 2 years indicate that patients:

- were better able to perform physical activities without limitations due to health
- had fewer limitations due to pain
- evaluated personal health higher
- experienced more energy or 'pep'
- performed more social activities without interference from physical or emotional problems, and
- had higher productivity at work, home, or school.

Importantly, a mean improvement of 10.8 from baseline of 30.9, or 35% in the leflunomide SF-36 PCS score (which includes changes in all 8 domains) would be considered to exceed the proposed definitions of MCID. Despite a baseline SF-36 MCS score of 48.5, which approximates the US norm of 50, the mean improvement of 4.7 over 24 months would approximate proposed definitions of MCID.

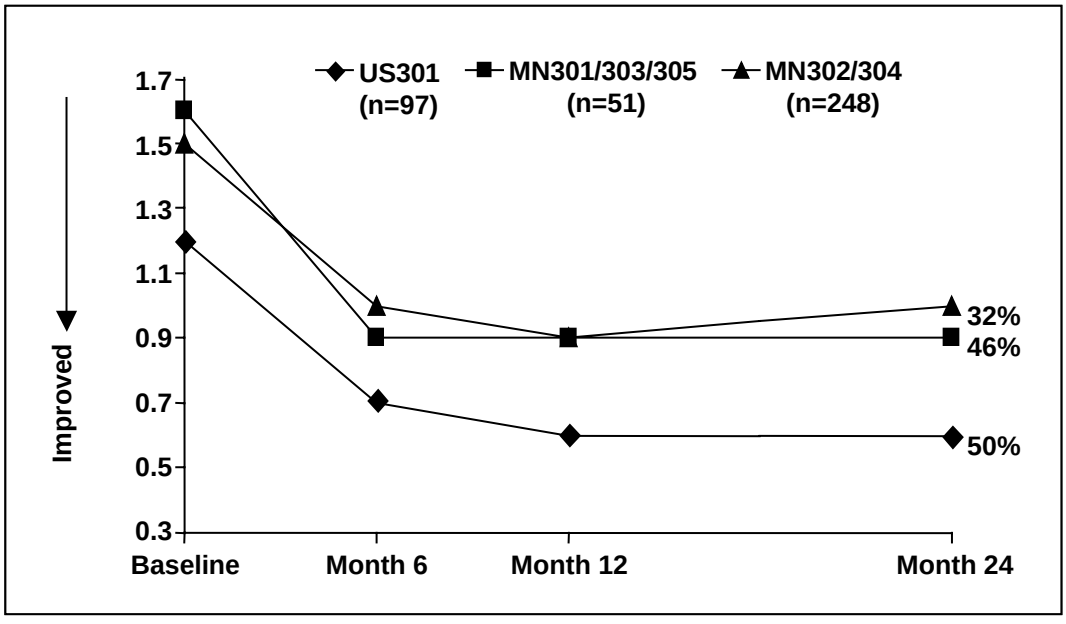
In summary, based on changes from baseline in mean scores of HAQ DI in all trials, as well as PET and SF-36 in US301, the majority of patients in the leflunomide treatment groups achieved improvements that would be clinically meaningful on an individual basis at 6 months and were still evident at 2 years.

Similar results from US301 can be seen with the PET top 5 score, ranging from 54% to 68% achieving MCID. Using a 5-point change for MCID in the SF-36 PCS and MCS summary scores, 69% of the year-2 cohort patients in the leflunomide treatment group achieved an MCID in the PCS at 6 months and 45% in the MCS at 6 months. These percentages were maintained after 2 years of treatment: 67% and 48% of patients maintained improvements in the PCS and MCS, respectively, greater than or equal to MCID of 5.

Change in HAQ DI over time with treatment

It has been suggested that RA patients treated under standard of care with DMARDs, NSAIDs and corticosteroids will worsen in HAQ DI by approximately 0.031 points per year, or 0.062 points over 2 years [17]. Several cohort studies have shown that HAQ DI scores remain stable in some patients but deteriorate in others over 2 to 5 years of conventional treatment with standard-of-care agents. According to Uhlig et al., after 2 years of treatment (52% of patients with DMARDs, 32% with corticosteroids, and 38% with NSAIDs), HAQ DI scores remained at the same level as baseline (0.9) [23]. Young et al reported that, in a cohort of 732 patients with early disease of whom 84% received DMARDs, 60% worsened in physical function over 5 years (by HAQ DI and ACR functional grade) [28]. In comparison, in the three phase III trials reported in this *Briefing Document*, patients receiving leflunomide had higher HAQ DI scores at baseline (US301: 1.2, MN301: 1.6, MN302: 1.5) yet showed improvements within or before 6 months treatment that were maintained over 2 years in patients continuing treatment (endpoints: US301: 0.6, MN305: 0.87, MN304: 1.02). These values reflect improvements of 0.60, 0.73, and 0.48, respectively, that all considerably exceed the MCID of 0.22.

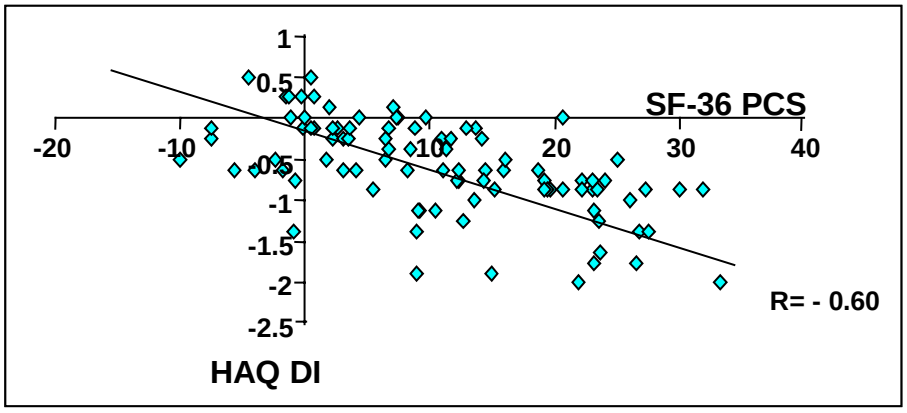
Leflunomide: Improvement in HAQ DI, Year-2 Cohorts, 0 to 24 months



Improvements in HAQ DI scores are reflected in SF-36

In US301, where both generic and disease-specific physical function and health-related quality of life instruments were utilized, improvements observed with leflunomide treatment in HAQ DI and individual scores were closely reflected by improvements in SF-36 (see figure below). These changes occurred not only in those domains of SF-36 directly associated with physical function (e.g., physical functioning, role physical and bodily pain) as might be expected, but also in domains such as vitality and role emotional. It should be noted that US301 was the first RA study that showed a change in SF-36, as opposed to an early RA study (MIRA) by Tuttleman et al. [22]. Although a generic measure of health-related quality of life was not included in the European studies MN305 and MN304 (partly because SF-36 translations were not available when the studies were initiated), improvements in HAQ scores and HAQ DI observed in the leflunomide treatment groups were of similar magnitude to those in US301. It can therefore be strongly expected that similar improvements would have been reflected in a generic measure of health-related quality of life in the two European leflunomide patient populations, both in magnitude of effect as well as in domains other than physical function.

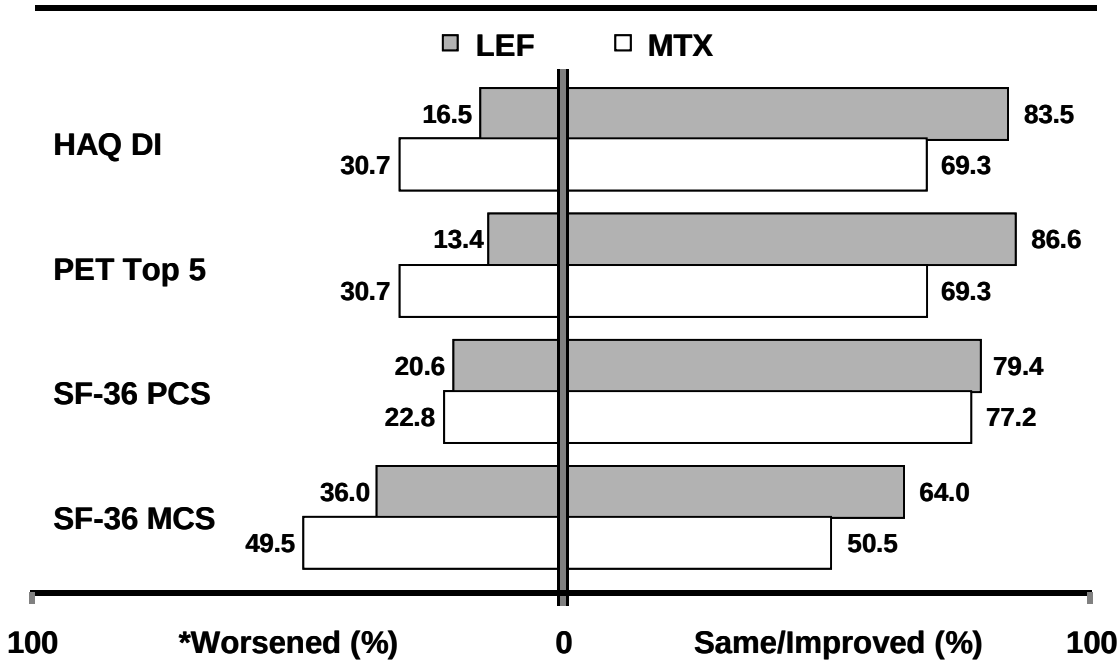
US301: Correlation HAQ DI and SF-36 PCS



Patient-reported improvements

Patient-reported responses for the HAQ DI, PET Top 5, SF-36 PCS, and SF-36 MCS are shown for study US301 were compared for leflunomide and methotrexate. As shown in the figure below, for all four measures, more patients treated with leflunomide than with methotrexate reported that they had improved or stayed the same from baseline to month 24, while more patients treated with methotrexate than with leflunomide reported that they had worsened during the same interval.

Responses in Patient-Reported Outcomes: US301 Year-2 Cohort



**** Any worsening from baseline to month 24***

One of the health-transition questions asked in the SF-36 is: "Compared to one year ago, how would you rate your health in general now?" For patients in study US301, 90% of leflunomide patients achieving MCID in the HAQ DI responded that they had improved, while of the patients who responded that they had improved, 72% had achieved MCID.

For the SF-36 PCS, 91% of leflunomide patients in study US301 who achieved MCID responded that they had improved, while of the patients who responded that they had improved, 61% had achieved MCID.

Number Needed to Treat

The Number Needed to Treat (NNT) approach to patient-reported outcomes evaluates the number of patients that need to be treated to obtain a clinically meaningful improvement, based on achieving MCID. The NNT for leflunomide compared to placebo and methotrexate demonstrate consistent results. These single-digit NNTs demonstrate the robustness of the improvements seen with leflunomide treatment.

Number Needed to Treat to Achieve MCID HAQ DI, PET Top 5, SF-36 at 12 Months

Category	Methotrexate *versus* Placebo	Leflunomide *versus* Placebo	Leflunomide *versus* Methotrexate
HAQ DI [MCID: -0.22]	5.6	2.9	6.5
PET Top 5 [MCID: +5.0]	13.6	4.3	6.6
SF-36 PCS [MCID: +5.0]	16.3	4.6	6.4

Source: Strand et al. *Arthritis &Rheumatism.* 2001;44:S187.

SF-36 summary scores (PCS and MCS)

The SF-36 PCS measures much more than decrements in physical function experienced by patients with active RA; it measures how these decrements affect the patients' day-to-day activities. Improvements in this summary score correlated with similar changes in physical function (by HAQ and PET) and ACR responder status [19]. They also correlated, on an individual patient basis, with ACR responses of ≥20%, ≥50%, and ≥70%.

Using data from the general US population provided in the SF-36 manual [25], PCS and MCS can be used to determine the impact of the level of improvement on a general population through a standard normative technique (also described in the manual). The following table shows an extrapolation of limitations associated with the PCS in the leflunomide treatment group in US301 at baseline (30.8) to the corresponding percentage of adults in the general US population with a similar PCS. Thus, a PCS of 30.8 in the general US population would reflect 44.5% reporting limitations with walking a block, 66.9% reporting limitations climbing stairs, 88.5% reporting difficulty at work, and 21.4% reporting very severe pain. Similarly, limitations associated with PCS in the leflunomide group at endpoint (41.7) are shown. This demonstrates that leflunomide administration could be expected to greatly reduce the percentage of adults who would report a limitation in performing daily physical activities.

Leflunomide: Percentages of the general US population reporting limitations corresponding to SF-36 PCS scores (US301)

Category	Associated baseline limitations	Associated endpoint limitations	Reduction in limitations
Walking a block	44.5%	17.3%	27.2%
Climbing stairs	66.9%	27.7%	39.2%
Difficulty at work	88.5%	47.2%	41.3%
Pain (very severe)	21.4%	4.1%	17.3%

The following table shows a similar extrapolation for the methotrexate group in US301. The PCS at baseline for methotrexate was 30.2 and endpoint was 38.8.

Methotrexate: Percentages of the general US population reporting limitations corresponding to SF-36 PCS scores (US301)

Category	Associated baseline limitations	Associated endpoint limitations	Reduction in limitations
Walking a block	44.5%	38.4%	6.1%
Climbing stairs	66.9%	43.2%	23.7%
Difficulty at work	88.5%	67.5%	21.0%
Pain (very severe)	21.4%	12.4%	9.0%

A comparison of the extrapolated data from the general US population reveals that the leflunomide group would have a greater reduction in limitations relative to the methotrexate group. Specifically, there would be 21.1% fewer patients with limitations in walking a block, 15.5% fewer with limitations in climbing stairs, 20.3% fewer with difficulties at work and 8.3% fewer with reports of very severe pain in the leflunomide group compared with methotrexate.

The following table shows an extrapolation of limitations associated with the MCS in the leflunomide treatment group in US301 at baseline (48.5) to the corresponding percentage of adults in the general US population with a similar MCS. Thus, a MCS of 48.5 in the general US population would reflect 38.3% reporting feeling happy, 21.0% reporting "lots of energy", 31.0% reporting accomplishing less at work, and 14.0% reporting social limitations. Similarly, limitations associated with the MCS in the leflunomide group at endpoint (53.1) are shown. Although the mean improvement in SF-36 MCS following leflunomide treatment in US301 was small (4.7 points), it nonetheless reflects important changes in a majority of patients in their ability to feel happy, accomplish more at work, and engage in more social activities with increased energy. SF-36 MCS for methotrexate would be expected to show similar results based on similar baseline (49.8) and endpoint (52.5) scores.

Leflunomide: Frequencies of the US general population reporting improvements corresponding to SF-36 MCS responses (US301)

Category	Associated baseline reporting	Associated endpoint reporting	Improvement in activities
Feeling happy	38.3%	63.7%	25.4%
Lot of energy	21.0%	36.3%	15.3%
Category	**Associated baseline limitations**	**Associated endpoint limitations**	**Reduction in limitations**
Accomplish less at work	31.0%	11.5%	19.5%
Social limitations	14.0%	8.2%	5.8%

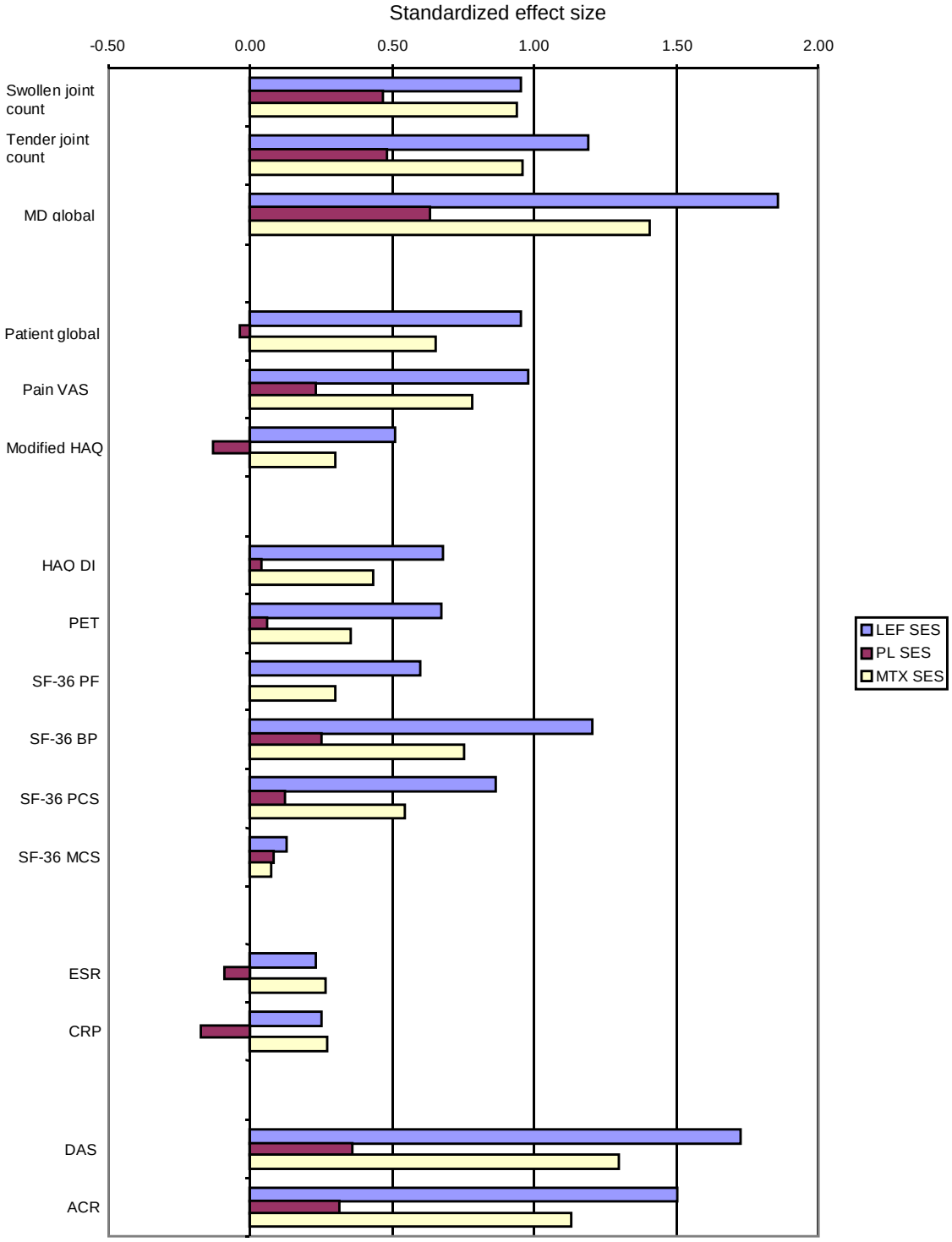
Using the extrapolated data, there are no differences between the methotrexate and leflunomide groups, however, both groups would show meaningful population improvements in the above MCS categories.

Patient-reported outcomes are more sensitive to active treatment effects

In the leflunomide group in US301, the relative efficiencies of HAQ, PET, and SF-36 to detect a treatment effect compared with the tender joint count were more sensitive than traditional measures [21]. The HAQ, PET, physical functioning domain, and PCS score of the SF-36 performed sufficiently well that clinically important differences could be detected with a statistical significance of <0.001.

As shown in the following figure, disease-specific (HAQ and PET) and generic (SF-36) patient-reported measures of physical function and health-related quality of life show large standardized effect sizes in all outcome measures for active treatment, and the smallest for placebo—indicating better discriminative ability to identify a true treatment effect [3]. Physician-reported measures demonstrate effect sizes that approximate a moderate effect in placebo patients. In contrast, patient reported measures demonstrate little to no change in placebo patients. The effect sizes for active therapies are relatively greater, suggesting that these patient-reported measures are more sensitive to true treatment effects, supporting an earlier publication by Tugwell et al. [21].

Patient-reported outcomes are not subject to placebo responses and are more sensitive to treatment effects



BP = bodily pain, CRP = C-reactive protein, DAS = disease activity score, ESR = erythrocyte sedimentation rate, PF = physical function, VAS = visual analog scale

Source: [3]

Summary of clinical relevance

Mean improvements in patient-reported assessments of physical function and health-related quality of life in the leflunomide treatment groups were consistent across all three Phase III trials and exceeded published values by two or more times, representing minimum clinically important differences (MCID in HAQ DI). Although several cohort studies have shown that HAQ DI scores remain stable or deteriorate in RA patients receiving standard of care, despite higher HAQ DI scores at baseline, the majority of patients receiving leflunomide treatment reported clinically important improvement at 6 months (57 to 71% patients) that remained evident at 24 months (60 to 70%).

Improvements reported with leflunomide treatment in the HAQ DI and its individual subscales were closely reflected by improvements in SF-36. Changes were reported not only in those SF-36 domains directly associated with physical function (e.g., physical functioning, role physical, and bodily pain) as might be expected, but also in vitality and role emotional. These patient-reported measures of physical function and health-related quality of life (HAQ DI, PET, and SF-36) appear most sensitive to change and not susceptible to a placebo response, thereby best reflecting a true treatment effect. These changes indicate reduced limitations in physical function and improved health-related quality of life, and are those outcomes that are most important to the patient.

1.4 Signs and Symptoms of RA over 2 years

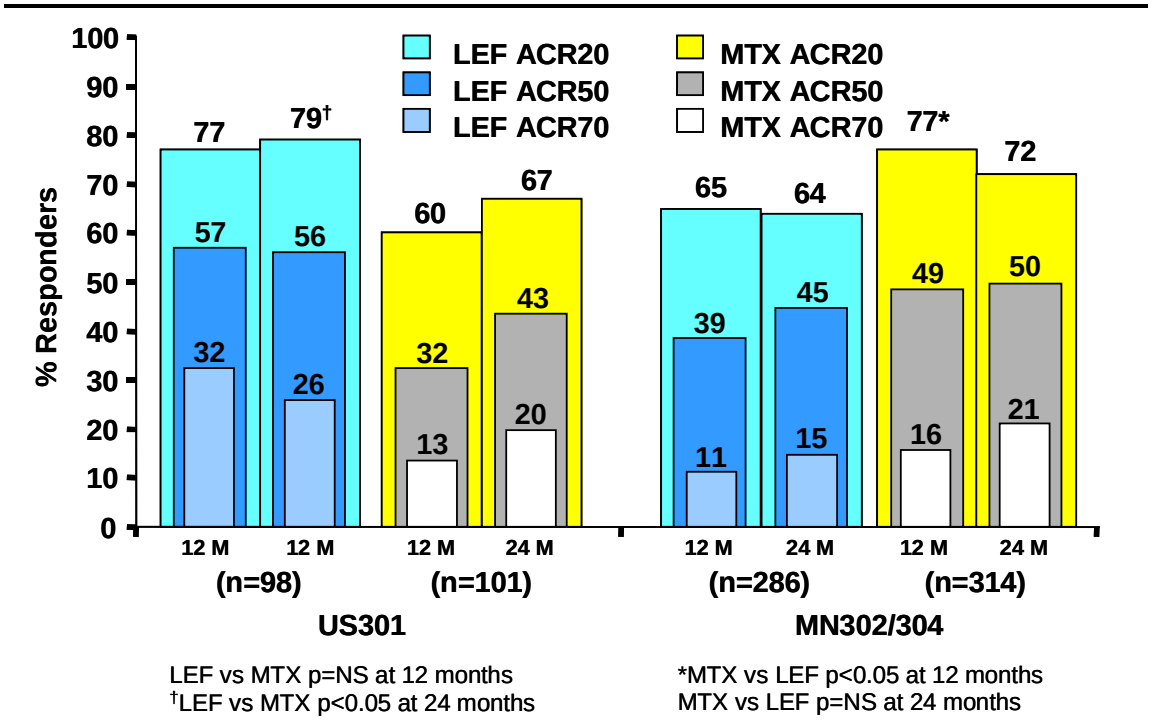
Improvements in signs and symptoms of RA as demonstrated by ACR response rates after 6 and 12 months were maintained over 2 years of leflunomide treatment. Improvement in the individual components of the ACR responder criteria were sustained over the second year of leflunomide treatment in all three Phase III studies.

Percentages of ACR responders in leflunomide year-2 cohorts of Phase III studies

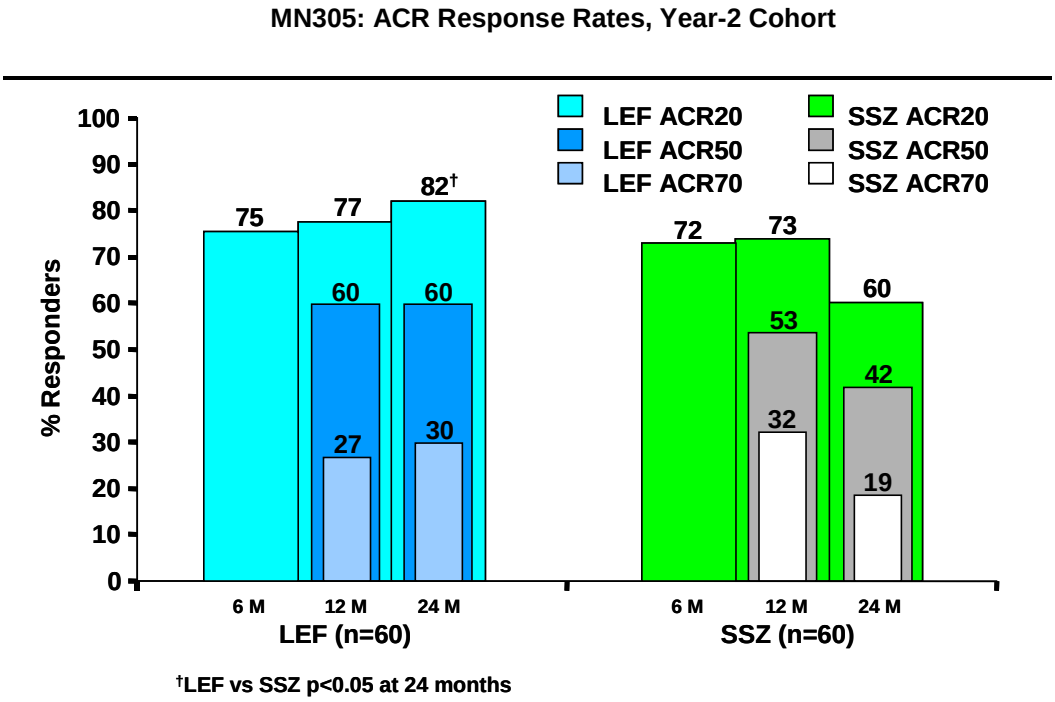
	ACR 20%		ACR 50%		ACR 70%	
Study	**Month 12**	**Month 24**	**Month 12**	**Month 24**	**Month 12**	**Month 24**
US301	77%	79%	57%	56%	32%	26%
MN305	77%	82%	60%	60%	27%	30%
MN304	65%	64%	39%	45%	11%	15%

The figure below shows a comparison of response rates for the ACR20, ACR50, and ACR70 at 12 and 24 months for year-2 cohort patients treated with leflunomide and methotrexate in studies US301 and MN302/304.

US301 and MN304: ACR Response Rates, Year-2 Cohort



The figure below shows a comparison of response rates for the ACR20, ACR50, and ACR70 at 12 and 24 months (and in addition, at 6 months for the ACR20 only) for year-2 cohort patients treated with leflunomide and sulfasalazine in study MN301/303/305.



Detailed descriptions of signs and symptom results may be found in the study synopses provided in *Appendix 2*.

1.5. Clinical Trial Safety Results over 2 years

Safety results in the Phase III studies are presented for two cohorts:

- **ITT cohort:** The cohort of patients initially enrolled into US301, MN301 and MN302 who took at least 1 dose of study medication.
- **Year-2 cohort:** The cohort of patients who continued into a second year of treatment.

Results for methotrexate-treated patients are presented separately for US301 and MN304. This is because folate was administered to 98% of patients in US301 as mandated by the protocol, while only 11% of methotrexate patients in study MN304 received folate, generally after an adverse event.

1.5.1 Exposure to study medication

Exposure to study medication in the Phase II and III studies is summarized in the following table.

Exposure to study medication in Phase II and III studies

	No. patients exposed				
Treatment group	**Total**	**≥6 months**	**≥12 months**	**≥18 months**	**≥24 months**
Leflunomide	1468	1070	970	709	558
Methotrexate	688	549	499	393	367
Sulfasalazine	133	76	69	55	48
Placebo	322	144	40	30	28

A total of 1468 patients were exposed to leflunomide in the Phase II and III studies, whereby 1070 patients were treated for at least 6 months and 970 patients for at least 12 months. Of the patients who completed 12 months of leflunomide treatment, 450 started a second year of treatment in the Phase III studies (i.e., in US301, MN305, or MN304) and were, therefore, included in the integrated analysis of safety over 2 years (year-2 cohort).

Exposure to study medication and reasons for withdrawal in the year-2 cohort of the Phase III studies are summarized by treatment group in the following table. The patients were treated for an average of 23 to 24 months in all treatment groups.

Exposure to study medication and reasons for withdrawal in Phase III studies: year-2 cohort

Parameter	Leflunomide (N=450)	Sulfasalazine (N=60)	MTX US301 (with folate) (N=101)	MTX 304 (without folate) (N=320)
Mean duration of treatment in months	23.7	23.0	22.8	23.7
Mean dose	19.5 mg/day	2.0 g/day	12.6 mg/week	12.2 mg/week
Median dose	20.0 mg/day	2.0 g/day	15.0 mg/week	10.0 mg/week
Withdrawals, no. (% patients)	59 (13.1)	13 (21.7)	21 (20.8)	43 (13.4)
Adverse event	23 (5.1)	8 (13.3)	8 (7.9)	15 (4.7)
Lack of efficacy	17 (3.8)	3 (5.0)	5 (5.0)	9 (2.8)
Other reason	19 (4.2)	2 (3.3)	8 (7.9)	19 (5.9)

1.5.2 Serious adverse events, deaths, and adverse events leading to withdrawal

Frequencies of serious adverse events, deaths, and adverse events leading to withdrawal in the year-2 cohort of the Phase III studies are given by treatment group in the following table.

Frequencies of serious adverse events and adverse events leading to withdrawal during year 2 in Phase III studies: year-2 cohort

	% patients			
Type of adverse event	Leflunomide (N=450)	Sulfasalazine (N=60)	MTX US301 (with folate) (N=101)	MTX 304 (without folate) (N=320)
Serious AEs	25.3	26.7	20.8	27.2
Related	3.1	8.3	2.0	1.6
Serious AEs, leading to withdrawal	0.9	5.0	6.9	1.6
Related	0.4	1.7	2.0	0.6
Death	0.7	0.0	1.0	2.2
Related	0.2	0.0	0.0	0.3
AEs leading to withdrawal	4.0	13.3	7.9	4.4
Related	2.4	10.0	4.0	3.4

Frequencies of serious adverse events were similar across treatment groups in the year-2 cohort but adverse events leading to withdrawal were less frequent in the leflunomide group than in the sulfasalazine and methotrexate groups.

As shown in the following table, the frequencies of serious adverse for the year-2 cohort were similar in the first and second years of leflunomide treatment and similar to those in the first year of treatment for the ITT cohort. As expected, serious adverse events leading to withdrawal were more common in the ITT cohort than in the year-2 cohort.

Frequencies of serious adverse events in leflunomide patients in Phase III studies: ITT and year-2 cohorts

	ITT cohort (N=824)		Year-2 cohort (N=450)			
	Onset in year 1		Onset in year 1		Onset in year 2	
Type of adverse event	n	%	n	%	n	%
Serious AEs	219	(26.6)	107	(23.8)	114	(25.3)
Related	45	(5.5)	11	(2.4)	14	(3.1)
Serious AEs, leading to withdrawal	50	(6.1)	–		4	(0.9)
Related	24	(2.9)	–		2	(0.4)
Death	6	(0.7)	–		3	(0.7)
Related	2	(0.2)	–		1	(0.2)

Three of the deaths in the leflunomide patients were assessed as possibly related to study medication:

- Patient 8/1004, MN302: This 74-year-old male stopped study medication (history of ulcerative colitis) and died suddenly a few days later due to respiratory and cardiac arrest.
- Patient 75/1004, MN302: This 34-year-old male died suddenly due to acute heart failure. The patient had a myocardial infarction with functional cardiac arrest prior to study entry.
- Patient 46/1002, MN304: This 57-year male died of esophageal carcinoma in the setting of Barrett's esophagus.

Frequencies of rare serious adverse events in the three Phase III studies are presented in the following table (rate/100 patient years).

Frequencies of rare serious adverse events in Phase III studies: ITT cohort over 2 years

	Rate/100 patient years			
Adverse event	Leflunomide (N=824)	Sulfasalazine (N=133)	MTX US301 (with folate) (N=190)	MTX 304 (without folate) (N=498)
Patient years	*1333*	*181*	*226*	*993*
Fatal infection	0.0	0.0	0.4	0.5
Sepsis, nonfatal	0.1	1.1	0.4	0.2
Malignancies	1.4	4.4	2.2	1.5
Lymphoproliferative disorders	0.2	1.1	0.0	0.1
Interstitial pneumonitis	0.0	0.0	0.4	0.4
Renal failure	0.0	0.0	0.4	0.2
Agranulocytosis	0.0	1.1	0.0	0.0
Vasculitis	0.8	0.6	0.4	0.5

There is no evidence that any of the above adverse events were more common with leflunomide than with sulfasalazine or methotrexate. There were no cases of fatal infection, interstitial pneumonia, renal failure, agranulocytosis, or pancytopenia in the 1333 patient years observed with leflunomide.

1.5.3 Most common adverse events

Frequencies of the most common adverse events in the ITT and year-2 cohorts of the Phase III studies are summarized in the following table.

Frequencies of adverse events that occurred in ≥10% of leflunomide patients in the ITT or year-2 cohort of the Phase III studies

	% patients		
	ITT cohort (N=824)	Year-2 cohort (N=450)	
Type of adverse event	Onset in year 1	Onset in year 1	Onset in year 2
All infections	49.2	56.0	50.7
Upper respiratory infection	25.8	30.7	27.1
Hypertension	10.2	10.7	9.6
Diarrhea	24.0	25.6	7.8
Nausea	12.9	10.0	2.4
Headache	11.2	11.3	3.8
Alopecia	13.7	15.3	2.9
Rash	11.7	12.2	7.1

The incidences of infections, respiratory infection and hypertension in the year-2 cohort were similar in years 1 and 2. Diarrhea, nausea, headache, alopecia, and rash are known to be associated with leflunomide treatment and occurred at much lower frequencies in the second year than the first year of treatment. Comparison of adverse event frequencies in year 1 revealed no relevant differences between the ITT and year-2 cohorts, thus showing that the incidences of adverse events during the second year of treatment in the year-2 cohort are representative of the total Phase III population.

1.5.4 Liver enzymes

Frequencies of alanine transaminase (ALT) and aspartate transaminase (AST) elevations in the ITT and year-2 cohorts of leflunomide patients are summarized in the following table.

Frequencies of ALT and AST elevations in leflunomide patients from the ITT and year-2 cohorts of the Phase III studies

Enzyme	No. (%) patients						
	ITT cohort (N=824)	Year-2 cohort (N=450)					
	Year 1 Occurrence		Year 1 Occurrence*		Year 2 Occurrence*		Year 2 elevation reversed to ≤2 x ULN after elevation
ALT							
>2 to ≤3 x ULN	38	(4.6)	23	(5.1)	13	(2.9)	10
>3 x ULN	25	(3.0)	11	(2.4)	8	(1.8)	6
AST							
>2 to ≤3 x ULN	25	(3.0)	11	(2.4)	4	(0.9)	4
>3 x ULN	10	(1.2)	5	(1.1)	9	(2.0)	8
Adverse events							
increased liver function tests	64	(7.8)	25	(5.6)	15	(3.3)	NA

ULN = upper limit of normal range, NA = not applicable
*In the year 2 cohort a subject with an elevation in year 1 and year 2 is counted twice.

The most sensitive enzyme to elevations, ALT, had a lower occurrence of elevations in year 2 than year 1 in the year-2 cohort. This was also true for adverse events of abnormal liver function tests and AST 2-3xULN. In most of these patients, values had reversed to below 2 x ULN by the end of the second year.

1.5.5 Hypertension

An analysis of hypertension is presented in the following table

Summary of hypertension in leflunomide patients from the ITT and year-2 cohorts of the Phase III studies

	No. (%) patients		
	ITT cohort (N=824)	Year-2 cohort (N=450)	
Criterion	Onset in year 1	Onset in year 1	Onset in year 2
Hypertension reported as AE	84 (10.2)	48 (10.7)	43 (9.6)
Total with hypertension at baseline	71	41	27
Diagnosis of hypertension at baseline	47	34	13
Blood pressure increased at baseline/screening	57	30	20
New-onset hypertension	13 (1.6)	7 (1.6)	16 (3.6)*
Systolic ≥160 mm Hg (at ≥2 visits)	13	7	14
Diastolic ≥90 mm Hg (at ≥2 visits)	8	3	11
Systolic ≥160 mm Hg and diastolic ≥90 mm Hg	8	3	10
Concomitant NSAIDs	13	7	12
Concomitant steroids	9	5	12
Mean change from baseline (mm Hg)			
Systolic blood pressure	0.6	–0.1	1.7
Diastolic blood pressure	1.0	1.0	0.5

* One of these patients also had an event during year 1, which resolved and recurred in year 2.

The incidence of hypertension reported as an adverse event decreased slightly during the second year of treatment with leflunomide compared with the first year (9.6% versus 10.7%). As in the first year most of the hypertension events were mild to moderate. All but 3 of the 43 patients with hypertension reported in the second year were treated with antihypertensive medications. One of the adverse events was a serious adverse event but no patient discontinued leflunomide treatment during the second year due to hypertension.

In the year-2 cohort, fewer patients had hypertension reported as an adverse event with onset in year 2 than in year 1. The incidence of new-onset hypertension (not present at baseline by history or blood pressure measurement) increased numerically from 1.6% (7 patients) during the first year of treatment with leflunomide to 3.6% (16 patients) during the second year, although this was not statistically significant (p=0.089). One of these patients had new-onset hypertension during the first year, which resolved and then recurred during the second year; therefore, 3.3% (15 patients) had a first occurrence of new-onset hypertension in year 2. The mean change in systolic blood pressure from baseline during the second year was a 1.7 mm Hg increase, compared to a 0.1 mm Hg decrease during the first year. The mean change in diastolic blood pressure from baseline during the second year was slight (0.5 mm Hg) and less than the change in year 1 (1.0 mm Hg).

It appears that leflunomide treatment has a mild effect on blood pressure. The contribution of concomitant NSAIDs and steroids to blood pressure changes cannot, however, be excluded.

1.5.6 Safety conclusions from clinical trials

In summary, the adverse event profile of leflunomide during the second year of treatment was similar to that during the first year of treatment and no new types of adverse events emerged. The incidences of diarrhea, nausea, headache, alopecia, rash, hypertension, and increased liver function tests decreased in the second year of treatment.

2. BENEFIT-RISK ANALYSIS OF LEFLUNOMIDE

2.1 Clinical and Post-Marketing Safety Data

The safety profile for ARAVA® is based on three types categories of safety information:

- Data from randomized, controlled, clinical trials
- Post-marketing safety surveillance data for ARAVA®
- Epidemiologic analysis of large cohorts of RA patients.

Data on safety and adverse events from clinical trials and post-marketing surveillance for ARAVA® are reviewed in the Aventis response to a March 28, 2002 Public Citizen Health Research Group petition to the FDA requesting withdrawal of ARAVA® (leflunomide) Tablets from the market. The Aventis response was submitted to the FDA on August 8, 2002 and is included in Appendix 1.

2.2 Epidemiologic Studies

2.2.1 Retrospective cohort study

The objective of this post-marketing, retrospective cohort study was to compare rates of adverse events (AEs) amongst leflunomide users to patients taking DMARDs (e.g., gold salts, azathioprine, hydroxychloroquine, D-penicillamine, sulfasalazine and the biologics etanercept and infliximab), alone and in combination.

This study relied on the 6.5 million-member claims database of Aetna, a US health insurer. Follow-up occurred from September 1998 through the end of December 2000. A diagnosis of Rheumatoid Arthritis and use of a DMARD were required for entry into the cohort. Subjects had to be 18 or over at time of entry. Exposure and time on drug was identified by dispensed prescription data. Outcomes included hepatic, hematologic, hypertensive, pancreatic, respiratory, and severe skin adverse events (AEs). Rates were computed using Poisson regression and were adjusted for age, sex, and comorbidities.

The study assembled more than 40,500 RA patients and 83,000 person-years (PY) of follow-up, making it the largest RA cohort study ever conducted. The leflunomide monotherapy exposure group had significantly fewer AEs than DMARD and MTX groups. The leflunomide group had rates of hepatic, hematologic, pancreatic, pneumonitis, and severe skin AEs that were comparable to DMARD and MTX. Leflunomide patients had significantly lower rates of hypertension and upper respiratory AEs compared to DMARD and MTX. The combination of leflunomide + MTX exposure group had AE rates that were comparable to other combination therapies. The exposure group no-DMARDs generally had the highest rates observed in this study for all AEs. This is likely due to a 'depletion of susceptibles' effect and channeling bias, in which patients who experience an AE on a drug will be taken off and put on another, less toxic regimen.

Although data on disease severity, OTC use, and history of RA were missing, it was clear that in this large population, leflunomide's safety profile is similar to that of other DMARDs.

The full report of this study is presented in *Appendix 4*.

2.2.2 Bi- cohort, nested, case-control study

The objective of this study was to replicate the retrospective (Aetna) cohort study using different databases and a slightly different design. This study relied upon the combined data from the Protocare claims database (10 million members) and the PharMetrics database (16 million members). Follow-up occurred from September 1998 through December 2001. Subjects were entered if they had an RA diagnosis, had a prescription for a DMARD after September 1998, were 18 or over at entry, and had not experienced one of the endpoints of interest in the 90 days prior to entry. Exposures included methotrexate, leflunomide, other DMARDs, and biologic DMARDs.

The combined databases had almost 42,000 persons who were prescribed a DMARD after September 1998 and a total of 51,315 person-years of follow-up time. Three-quarters of the cohort were women. The average age of Protocare subjects was 59, compared to 49 for PharMetrics subjects. There were 90 events per 10,000 PY for all events of interest combined, and 5 per 10,000 PY for severe hepatic events, 27 per 10,000 PY for hematologic events, 16 per 10,000 PY for pancreatitis, 42 per 10,000 for opportunistic infections and sepsis, less than 1 per 10,000 PY for severe skin disease, 2 per 10,000 PY for pneumonitis, and 1 per 10,000 PY for lymphoma. Using methotrexate as the reference, the adjusted rate ratios for leflunomide were not significantly different from 1 for any serious adverse event (RR =1.1), serious hepatic events (RR = 0.9), serious hematologic events (RR = 0.8), serious pancreatitis events (RR = 1.5), and serious opportunistic infections and septicemia events (RR = 0.9). There were too few events for rate calculations of severe skin, pneumonitis, and lymphoma events. Of note were the generally elevated RRs for the biologic DMARDs, especially for any event, serious liver events, and opportunistic infections and septicemia events.

This study affirms the earlier Aetna cohort study in that adverse events amongst leflunomide patients do not occur more often than they do in methotrexate patients.

The full report of this study is presented in *Appendix 5*.

2.2.3 Proportional reporting ratio analysis

The objective of the proportional reporting ratio (PRR) analysis was to determine if reports of adverse events amongst leflunomide users are inconsistent with similar reports amongst other DMARD users. PRR is a signal-generating tool, and is not used to confirm hypotheses. Proportional reporting ratio analysis compares spontaneous reports of suspected adverse reactions of different drugs where the true number of patients exposed to a drug is unknown.

PRR analysis is a useful statistical tool, widely employed by the Medicines Control Agency (MCA) in the UK. It is calculated using a 2 x 2 table, as follows:

	reaction of interest	all other reactions
drug of interest	***a***	***b***
all other drugs	***c***	***d***

PRR is calculated as a/(a + b) divided by c/(c + d). Criteria to evaluate the PRR come from several sources and are similar: a minimum of three reported cases are needed; a PRR of at least 3 and an associated X^2 over 5 or a PRR of at least 2 and an associated X^2 over 4 are considered possible signals. The data used are limited in that there is no way to assess the indication for a particular drug, so in the situation where a specific drug is used for more than one condition (e.g., as is the case with methotrexate), there is no way of adjusting for potential confounding by indication. The calculated PRR used the entire database of the FDA as a comparison (results which were not different than when DMARDs were used as a comparison group).

The results showed that specific AE reports of leflunomide, as a proportion of all leflunomide reports was not different than other drugs, with the possible exceptions of interstitial lung disease PRR and vasculitis PRR. These signals have been further examined using epidemiologic data and have been found to be unsupported.

The full report of this study is presented in *Appendix 6*.

2.2.4 Reporting rate analyses

The objective of this analysis was to examine the comparative reporting rates of various AEs of leflunomide and other DMARDs. Since this method relies on spontaneous report data, it is used for signal generation.

Spontaneous reports (numerator data) were obtained from the FDA via QScan, a commercial software vendor that offers access to the more than two million adverse event cases reported to the FDA made available through the Freedom of Information Act, using proprietary mapping tools and techniques. Denominators (sales data) were obtained from IMS and converted into person-year exposures. Leflunomide, methotrexate, infliximab, and etanercept were compared. Adverse events of interest included hepatic failure, interstitial lung disease, tuberculosis and sepsis, bullous conditions, lymphoma, demyelinating disorders, hypertension, vasculitis, and pancytopenia.

Using this method to evaluate potential signals from spontaneous reports, none were found for leflunomide. Spontaneous report analysis is made difficult by under-reporting, the Weber effect (i.e., reports are more frequent closer to time of launch and for a period of about two years, then drop off substantially), lack of interest by professionals to report, potential confounding by indication (i.e., the AE is caused by the condition being treated, not the drug), and poor quality reporting data. Compared to the two biologic DMARDs, which were launched approximately the same time as leflunomide, there does not appear to be any signals.

Using this method of analysis, the AE profile of leflunomide appears comparable to that of biologic DMARDs, with lower rates for certain events, the full report of this study is presented in *Appendix 6*.

2.2.5 Meta-analysis

The objective of this study was to compare the rates of adverse events seen in phase III clinical trials; specifically, leflunomide was compared to methotrexate and to sulfasalazine.

Adverse event rates were cumulated from clinical trials US301 (placebo-controlled trial of leflunomide versus methotrexate), MN301/303/305 (placebo-controlled trial and extensions of

leflunomide versus sulfasalazine), and MN302/304 (leflunomide versus methotrexate). The rates are presented on a L'Abbé scatter plot (line-of-identity graph) for ease and sensibility of interpretation.

The results of this meta-analysis show that Serious and Serious and Related adverse events all occur more often amongst the methotrexate and sulfasalazine users. Methotrexate and sulfasalazine also had higher rates of pain, blood, and cardiovascular AEs. Skin (rash) and hypertension occurred more often amongst leflunomide users. Leflunomide had higher rates of infection and abnormal liver tests compared to sulfasalazine, and lower rates compared to methotrexate.

Using L'Abbé scatter plots to assess the rates of AEs reported in clinical trials of leflunomide, the two comparator agents (methotrexate and sulfasalazine) had higher rates of Serious and Serious and Related events, as well as higher rates of cardiovascular, blood, and pain AEs. Leflunomide had higher rates of skin rash and hypertension.

The full report of this study is presented in *Appendix 7*.

2.2.6 Liver transplant analysis

The objective of this study was to determine how many liver transplant cases have been reported in which leflunomide or methotrexate is listed as the etiology.

Data were requested by and received from the United Network for Organ Sharing (UNOS). UNOS administers the nation's only Organ Procurement and Transplantation Network (OPTN), established by the US Congress in 1984. Through the OPTN UNOS collects and manages data about every transplant event occurring in the United States; facilitates organ matching and placement processes; and helps to develop organ transplantation policy. All data on liver transplants from 1 January 1998 through 31 July 2002 were reviewed for drug involvement of either methotrexate or leflunomide.

In 1998, three liver transplants listed methotrexate hepatotoxicity as a cause or diagnosis; in 1999 and 2000, one transplant each year listed methotrexate; in 2001, six transplants listed methotrexate; and through 31 July 2002, four cases listed methotrexate toxicity as a reason for the procedure. In that same time period, no cases listed leflunomide.

Based on a review of the UNOS liver transplant data, methotrexate toxicity was listed as the diagnosis for 15 liver transplants from January 1998 through July 2002. In that same period, leflunomide toxicity was not listed as the diagnosis for liver transplant.

2.2.7 National Data Bank for Rheumatic Diseases

Data from the National Data Bank for Rheumatic Diseases regarding rates of serious liver toxicity in patients taking leflunomide or methotrexate were published in abstract form and presented by Dr. Fred Wolfe at the American College of Rheumatology 2002 Annual Scientific Meeting. He reported that the rates were low and that there was no significant difference between leflunomide and methotrexate in the percent of patients with self-reported liver adverse events or in rates of liver adverse events per 100 patient-years [39].

Treatment effectiveness in the community has been evaluated by Dr Fred Wolfe based on data from the National Data Bank for Rheumatic Diseases and reported at the American College of Rheumatology 2001 and 2002 Annual Scientific Meetings. Data were evaluated using the time

patients remain on treatment [40] and also by an expanded definition of treatment failure, i.e., time to treatment discontinuation or addition of another DMARD [41]. In both of these measures of treatment effectiveness in the community, leflunomide and methotrexate were comparable.

2.3 Benefit-Risk Conclusions

The accepted standard of care for patients with RA is aggressive, early treatment with DMARDs, which slow and potentially alter the course of the disease. However, no single DMARD is effective in all patients, and secondary failures (loss of efficacy) are not uncommon. Accordingly, most patients with active RA require the progressive addition or change of treatments over time. Serious and sometimes fatal adverse events have been reported with each of these therapies. The need for alternative therapies remains the driving force behind recent development and approval of new treatments for RA over the last four years and a half years.

Several epidemiological approaches were used to compare the adverse event profile of leflunomide with those of other DMARD's and have been described in Section 2.2. While none of the approaches are without limitations, the results of all analyses taken together show an adverse event profile for leflunomide comparable to methotrexate and other DMARDs.

The efficacy and safety data confirm that ARAVA® is an important advance in the treatment of RA and should remain available to the many thousands of individuals who benefit from the use of the drug. The chronic, progressive, and destructive nature of RA warrants the use of DMARDS early in the disease process. ARAVA® has been clinically proven to have efficacy in early and advanced disease, with rapid onset of therapeutic effect and sustained benefit during long-term therapy.

These established benefits must be weighed against its recognized risks, in the context of other available therapies and the severity of the disease. The risk of serious and sometimes fatal adverse events has, unfortunately, been observed with most prescription medications – and all DMARDs, including ARAVA®. Specifically, the safety data from randomized controlled trials show the overall percentage of patients with adverse events who were treated with ARAVA® was generally comparable to that of patients who received methotrexate and sulfasalazine. Importantly, nothing in the post-marketing experience changes the acceptable benefit-risk profile established by the controlled clinical studies.

When weighed against the benefits of the drug, its impact on the disease course, and the limitations of other available therapies, the risks of ARAVA® treatment are clearly outweighed by its substantial benefits.

3. OVERALL CONCLUSIONS

At 12 months in the ITT population, leflunomide was statistically significantly superior to placebo in improving physical function and health-related quality of life as assessed by the HAQ DI, and these changes were clinically significant [25]. Superiority to placebo was demonstrated consistently across all eight HAQ DI subscales in both placebo-controlled studies (US301 and MN302).

The SF-36 further addresses physical function as well as social and emotional function. In US301 at 12 months in the ITT population, leflunomide treatment resulted in statistically significant improvements compared to placebo in 5 of the 8 SF-36 scales (physical functioning, pain, general health perception, vitality, and social functioning), the SF-36 Physical Component Summary score, and the Work Productivity Questionnaire.

The improvements in physical function and health-related quality of life demonstrated at 6 and 12 months were maintained over 2 years. In patients continuing leflunomide for a second year of double-blind treatment in US301, MN305 (extension of MN301/303) or MN304 (extension of MN302), marked, clinically meaningful improvements from baseline in the HAQ DI were still observed at 24 months in all three trials, with no clinically meaningful differences between month 12 and month 24.

An ITT analysis over 24 months further showed leflunomide therapy to be statistically significantly superior to placebo on the HAQ DI in the two placebo-controlled trials, and on the SF-36 PCS in US301. Sensitivity analysis showed these results to be robust.

The adverse event profile of leflunomide during the second year of treatment was similar to that during the first year of treatment and no new types of adverse events emerged. The incidences of diarrhea, nausea, headache, alopecia, rash, hypertension and increased liver function tests decreased in the second year of treatment.

The rates of adverse events seen in these phase III clinical trials with leflunomide were compared to methotrexate and sulfasalazine in a meta-analysis. The results show that serious and serious and related adverse events all occur more often amongst the methotrexate and sulfasalazine users.

In addition to this meta–analysis several epidemiological approaches were used to compare the adverse event profile of leflunomide with those of other DMARD's. While none of the approaches are without limitations, the results of all analyses taken together show an adverse event profile for leflunomide comparable to methotrexate and other DMARDs.

In conclusion, these data and analyses support the efficacy of leflunomide with regard to improvement in physical function and health-related quality of life. Improvement in signs and symptoms was proven previously in the NDA, and continued improvement over 2 years has been demonstrated in these studies.

The safety of leflunomide has been demonstrated to be comparable to other widely used DMARD's.

4. LIST OF REFERENCES

1 Bombardier C, et al. Efficacy of cyclosporin A in rheumatoid arthritis: long term follow-up data and the effect on quality of life. Scandinavian J Rheumatol 1992;21(95):29-33.

2 Buchbinder R, et al. Which Outcome Measures Should be Used in Rheumatoid Arthritis Clinical Trials? Arth Rheum 1995;38(11):1568-1580.

3 Cohen S, Smolen J, Tugwell P, et al. Patient-based functional measures of drug treatment in rheumatoid arthritis trials are not susceptible to placebo effect. Arth Rheum 2000;43:S288.

4 Food and Drug Administration, Center for Drug Evaluation and Research. Guidance for industry: Clinical development programs for drugs, devices and biological products for the treatment of rheumatoid arthritis (RA). Rockville MD: February 1999.

5 Fries J, Spitz P, Kraines G, Holman H. Measurement of patient outcome in arthritis. Arth Rheum 1980;23(2):137-45.

6 Goldsmith C, Boers M, Bombardier C, Tugwell P for the OMERACT Committee: Criteria for clinically important changes in outcomes: development, scoring and evaluation of rheumatoid arthritis patient and trial profiles. J Rheumatol 1993;20:561-565.

7 Guzman J, Maetzel A, Peloso P, Yeung M, Bombardier C. Disability scores in DMARD trials: what is a clinically important change? Arth Rheum 1996;39:S208.

8 Kavanaugh A, Lipsky P, Furst D, Weisman M, St Clair W, Smolen J, et al. Infliximab improves long term quality of life and functional status in patients with rheumatoid arthritis. Arth Rheum 2000;3:147.

9 Kosinski M, Zhao SZ, Dedhiya S, Osterhaus JT, Ware JE. Determining minimally important changes in generic and disease-specific health related quality of life questionnaires in clinical trials of rheumatoid arthritis. Arth Rheum 2000;43:1478-87.

10 Kosinski M, et al. Determining clinically meaningful improvement in SF-36 scale scores for treatment studies in rheumatoid arthritis. Arth Rheum 2000;43:S140.

11 Lerner DJ, Amick III BC, Malspeis S, Rogers WH. A national survey of health-related work limitations among employed persons in the US. Disability Rehabilitation 2000;22(5):225–232.

12 Medeiros MMDC, Ferraz MB, Quaresma MR. The effect of rheumatoid arthritis on the quality of life of primary caregivers. J Rheumatol 2000;27:76-83.

13 Ramey D, Raynauld JP, Fries J. The Health Assessment Questionnaire 1992: Status and Review. Arthritis Care Res l992; 5(3):119-q 29.

14 Redelmeier DA, Lorig K. Assessing the clinical importance of symptomatic improvements. An illustration in rheumatology. Arch Intern Med 1993;153:1337-42.

15 Ruta DA, Hurst NP, Kind P, et al. Measuring health status in British patients with rheumatoid arthritis: reliability, validity and responsiveness of the short form 36-item health survey (SF-36). Brit J Rheumatol 1998; 37: 425-36.

16 Samsa G, Edelman D, Rothman M, et al. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. Pharmacoeconomics 1999;15(2):141-155.

17 Scott D, Pugner K, Kaarela K, et al. The links between joint damage and disability in rheumatoid arthritis. Rheumatol 2000;39:122-132.

18 Stewart A, Ware J. Measuring Function and Well-Being: The Medical Outcomes Study Approach. Duke University Press, Durham, NC, 1992.

19 Strand V, Tugwell P, Bombardier C, et al on behalf of the Leflunomide Rheumatoid Arthritis Group. 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. Arthr Rheum 1999;42:1870-8 1999.

20 Thumboo J, Fong KY, et al. Validation of the MOS SF-36 for quality of life assessment of patients with systemic lupus erythematosus in Singapore. J Rheumatol 1999;26(1):97-102.

21 Tugwell P, Wells G, Strand V, et al. Clinical improvement as reflected in measures of function and health-related quality of life following treatment with leflunomide compared with methotrexate in patients with rheumatoid arthritis: sensitivity and relative efficiency to detect a treatment effect in a 12-month placebo-controlled trial comparing leflunomide with methotrexate. Arth Rheum 2000; 43:506-514.

22 Tuttleman M, Pillemer S, Tilley B et al. A cross-sectional assessment of health status instruments in patients with rheumatoid arthritis participating in a clinical trial. J Rheumatol 1997; 24 (10): 1910-15.

23 Uhlig T, Smedstad L, Vaglum P, et al. The course of rheumatoid arthritis and predictors of psychological, physical and radiographic outcome after 5 years of follow-up. Rheumatol 2000;39:732-741.

24 Verhoeven AC, Boers M, van der Linden S: Validity of the MACTAR questionnaire as a functional index in a rheumatoid arthritis clinical trial. J Rheumatol 2000; 27:2801-9.

25 Ware JE, Kosinski M, Keller SD. SF-36 Physical and Mental Health Summary Scales: A User's Manual. The Health Institute, New England Medical Center, Boston, MA, December, 1994.

26 Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36):I. Conceptual framework and item selection. Medical Care 1992;30(6);473-483.

27 Wells G, Tugwell P, Kraag G, Baker P, Groh J, Redelmeier D. Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. J Rheumatol 1993;20:557-60.

28 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 (ERAS). Rheumatol 2000;39:603-11.

29 World Medical Association Declaration of Helsinki, Edinburgh, Scotland, October 2000.

30 Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD: Linking clinical relevance and statistical significance in evaluating intra-individual changes in HRQOL: Medical Care 1999; 37:469-78.

31 Wyrwich KW, Tierney WM, Wolinsky FD: Further evidence supporting an SEM based criterion for identifying meaningful intra-individual changes in HRQOL: J Clin Epidemiol 1999; 52:861-73.

32 Strand V, Cannon G, Cohen S, Ware J et al: Correlation of HAQ with SF-36: Comparison of Leflunomide to Methotrexate in patients with active RA. Arth Rheum 2001; 44:S187.

33 Strand V, Bombardier C, Maetzel A, Scott D, Crawford B: Use of minimum clinically important differences [MCID] in evaluating patient responses to treatment of RA. Arth Rheum 2001; 44:S187.

34 Zhao SZ, McMillen JI, Markenson JA, Dedhiya SD, Zhao WW, Osterhaus JT, Yu SS: Evaluation of the functional status aspects of health related quality of life of patients with osteoarthritis. Pharmacotherapy 1999; 19:1269-1278.

35 Ehrich EW, Bolognese JA, Kong S, Watson DJ, Zeng K, Seidenberg BC: Improvements in SF-36 mental health domains with treatment of OA result of decreased pain and disability or independent mechanism? Arth Rheum 1998; 41:S221.

36 Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N: Minimum perceptible clinical improvement with the WOMAC and global assessments in patients with osteoarthritis. J Rheumatol 2000; 27:2635-41.

37 Angst F, Aeschlimann A, Stucki G: Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 Quality of life measurement instruments in patients with osteoarthritis of the lower extremities. Arth Care Res 2001: 45:384-391.

38 Samsa G, Edelman D, Rothman M, et al. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. Pharmacoeconomics 1999; 15:141-155.

39 Wolfe F. Low rates of serious liver toxicity to leflunomide (LEF) and methotrexate (MTX): A longitudinal surveillance study of 14,997 LEF and MTX exposures in RA. Arthritis Rheum 2002; 46(9S): S375.

40 Wolfe F, et al. RA patients remain on leflunomide as long as methotrexate in clinical practice: Accumulating evidence of long-term tolerability and effectiveness of leflunomide in RA. Arthritis Rheum 2001; 44(9S): S144.

41 Wolfe F, et al. Toward a definition and method of assessment of treatment failure and treatment effectiveness: The case of leflunomide versus methotrexate. Arthritis Rheum 2002; 46(9S): S571

Docket No. 02P-0139

RESPONSE OF AVENTIS PHARMACEUTICALS INC. TO PUBLIC CITIZEN HEALTH RESEARCH GROUP'S PETITION REGARDING ARAVA® (LEFLUNOMIDE) TABLETS

Submitted By Aventis Pharmaceuticals Inc.
August 8, 2002

TABLE OF CONTENTS

Executive Summary And Introduction

Aventis Pharmaceuticals Inc. ("Aventis") submits this response to Public Citizen Health Research Group's ("HRG") March 28, 2002 Petition (the "Petition") requesting withdrawal of Arava® (leflunomide) Tablets from the market.[1] For the reasons discussed below, the Petition should be denied. The substantial and unique benefits of Arava® in the treatment of rheumatoid arthritis ("RA") clearly outweigh the risk of serious adverse events that may be associated with its use.

The Petition Mischaracterizes The Data

The Petition does not present a reasoned, scientific analysis; rather, it asserts unsubstantiated conclusions that ignore publicly available data, published literature, and standard medical practice.[2] In particular, the Petition trivializes the severity of RA, mischaracterizes the current clinical standard for RA patient care, distorts the safety and efficacy of alternative treatments, and ignores the treatment approaches recommended by the American College of Rheumatology ("ACR"). In short, HRG presents no substantive benefit-risk analysis for Arava® -- whether alone, in comparison to alternative therapies, or relative to the increased morbidity and mortality associated with RA. As stated by Gary S. Firestein, M.D., Chair of the FDA's Arthritis Advisory Committee, in his unsolicited letter opposing the Petition: "[M]erely describing the potential toxicity of an agent in a vacuum is not only insufficient but can be misleading." *See* Appendix A, 6/10/02 Correspondence from Dr. Firestein to FDA.

The Positive Benefit-Risk Profile Of Arava®

A substantive benefit-risk analysis requires objective scientific consideration of multiple factors. Among other things, one must consider the nature and severity of RA, the current standard of medical care and knowledge, the risks and benefits of other available therapies, and the efficacy and safety data associated with Arava®. A fair and balanced evaluation of the facts confirms the positive benefit-risk profile of Arava® -- and the continuing need for Arava® as an important treatment option for the many patients who suffer from this chronic, debilitating disease:

---

1 Aventis will refer to Arava® and leflunomide interchangeably throughout this Response

2 As demonstrated herein, HRG's certification that the Petition includes representative data unfavorable to its position is a misrepresentation. Substantial publicly available data that contradicts HRG's position has either been ignored or mischaracterized.

1. Nature And Severity Of Rheumatoid Arthritis

Rheumatoid arthritis is a debilitating autoimmune disease that affects more than 2 million Americans. The cause is unknown, and there is no known cure. Most patients exhibit a chronic fluctuating course of disease that can result in progressive joint destruction, deformity, disability, and, sometimes, premature death.[3] RA also affects other tissues and organs and results in more than 9 million physician visits and 250,000 hospitalizations per year.[4] It frequently affects patients in their most productive years, and the disability associated with the disease results in major economic loss to the individual and society.[5]

2. Standard Of Care And The Limitations Of Other Available Therapies

The accepted standard of care for patients with RA is aggressive, early treatment with disease-modifying antirheumatic drugs ("DMARDs"), which slow and potentially alter the course of the disease.[6] However, no single DMARD is effective in all patients, and secondary failures (loss of efficacy) are not uncommon. Accordingly, most patients with active RA require the progressive addition or change of treatments over time. Each of these therapies has been associated with serious and sometimes fatal adverse events. The need for alternative therapies remains the force driving recent development and approval of new treatments for RA in the last 4 years.

3. Clinically Proven Efficacy Of Arava®

Arava® has been shown in randomized, controlled trials to: (i) reduce the signs and symptoms of active RA; (ii) retard structural joint damage measured by radiographs; and (iii) improve physical function and health related quality of life. Arava® targets the underlying inflammatory process -- rather than just treating symptoms -- by inhibiting multiplication of T-cells believed to perpetuate the autoimmune response in RA. It is also effective in treating both early and long-standing disease, as long- and short-term therapy, and regardless of disease severity or previous exposure to other DMARDs. In clinical trials, Arava® had a faster onset of

---

3 Guidelines for the Management of Rheumatoid Arthritis, American College of Rheumatology Ad Hoc Committee On Clinical Guidelines (hereinafter "ACR Guidelines Update 2002"). *Arthritis Rheum*, 2002: 36(2) 328-46; Wolfe F. The burden of rheumatoid arthritis. *Am J Manag Care*. 1999; 5.S852-S859

4 ACR Guidelines Update 2002; Gabriel SE, Crowson CS, O'Fallon WM: The epidemiology of rheumatoid arthritis in Rochester, Minnesota, 1955-1985. *Arth Rheum* 1999; 42.415-20, Gabriel, S. E. The epidemiology of rheumatoid arthritis. *Rheum Dis Clin North Am.* 2001;27·269-81, Doran MF, Pond GR, Crowson CS, O'Fallon WM, Gabriel SE: Trends in incidence and mortality in rheumatoid arthritis in Rochester, Minnesota, over a 40 year period. *Arthritis Rheum* 2002; 46:625-31.

5 *Id*

6 ACR Guidelines Update 2002

action and equivalent improvement in physical function and radiographic progression when compared with methotrexate, the primary comparator drug referenced by HRG. Overall, clinical trial results confirm that the efficacy of Arava® is equivalent to both methotrexate and sulfasalazine.

Because of its unique properties and the need for additional DMARD treatments, Arava® received priority review by the U.S. Food and Drug Administration ("FDA"),[7] and an expert Arthritis Advisory Committee convened by the FDA unanimously supported marketing approval -- based on the same clinical trials referred to by HRG. Since the New Drug Application ("NDA") was approved in 1998, Arava® has been prescribed to more than 500,000 patients worldwide.

4. The Risk Of Adverse Events Is Outweighed By The Benefits Associated With Arava® Therapy

Throughout the Petition, HRG mischaracterizes and selectively cites clinical trial and post-marketing data, while ignoring critical information that clearly undermines its position. HRG's superficial analysis is not a substitute for a careful and thorough benefit-risk evaluation.

The facts confirm that Arava® is an important advance in the treatment of RA and should remain available to the many thousands of individuals who benefit from use of drug. The chronic, progressive, and destructive nature of RA warrants the use of DMARDs early in the disease process. Arava® has been clinically proven to have efficacy in early and advanced disease, with rapid onset of therapeutic effect and sustained benefit during long-term therapy.

These established benefits of Arava® must be weighed against its recognized risks, in the context of other available therapies and the severity of the disease. The risk of serious[8] and sometimes fatal adverse events has, unfortunately, been observed with most prescription medications -- and all DMARDs, including Arava®. In fact, treatment with each available DMARD has been associated with serious adverse events. None has a safety profile clinically proven to be superior to Arava®. Specifically, the safety data from randomized controlled trials show the overall percentage of patients with adverse events who were treated with Arava® was

---

7 A new drug application may receive priority review if "[t]he drug product, if approved, would be a significant improvement compared to marketed products . . . in the treatment, diagnosis, or prevention of a disease." Center for Drug Evaluation and Research, Manual of Policies and Procedures 6020.3. Priority review does not mean that less data is required to receive approval, but that, by regulation, the FDA will act on an expedited track due to the important therapeutic potential offered by the product.

8 The term "serious adverse events" is defined by the Code of Federal Regulations *See infra* footnote 94.

generally comparable to that of patients who received methotrexate and sulfasalazine. Importantly, nothing in the post-marketing experience changes the acceptable benefit-risk profile established by the controlled clinical studies.[9]

When weighed against the benefits of the drug, its impact on the disease course, and the limitations of other available therapies, the risks of Arava® treatment are clearly outweighed by its substantial benefits. Accordingly, the FDA and other regulatory bodies have correctly concluded that Arava® is one of the safe and effective therapies in the limited arsenal available to treat RA. Recently, the Agency for the Evaluation of Medicinal Products ("EMEA") and the Committee of Proprietary Medicinal Products ("CPMP") in Europe completed an exhaustive analysis of Arava® -- including the post-marketing and clinical trial data -- and concluded that "[t]he current benefit-risk assessment of ARAVA is positive and no change in SPC is needed" at the present time.[10] These conclusions continue to be supported by new data, including recent post-marketing clinical studies and surveillance reports, as well as the largest database analysis in RA patients. *See infra* subsection IV.B. The Petition does not support a contrary conclusion and, accordingly, should be denied.

* * * * *

The balance of this Response will address these matters in greater detail.[11] Part I will discuss the nature and severity of RA and known limitations of the available therapies. Part II describes the proven clinical efficacy of Arava®. Part III reviews the clinical trial safety data and post-marketing surveillance relating to Arava®. Part IV confirms the positive benefit-risk profile of Arava®. Finally, Part V demonstrates that the legal standard applicable to withdrawal of an NDA has not been satisfied and that the Petition should be denied.

---

9 When the post-marketing Arava® experience has produced an indication that there may be events not seen in clinical trials (or an increased frequency of previously observed events), Aventis has worked with the FDA to update the prescribing information and to notify physicians Aventis is currently working with the FDA to further update the prescribing information based on post-marketing information.

10 The EMEA performs administrative oversight and mobilizes scientific resources throughout Europe for medicinal products marketed in the European Union. The CPMP is a scientific standing Committee that evaluates medicinal products on behalf of the member states of the European Union and advises the European Union The SPC is the Summary of Product Characteristics, which is analogous to prescribing information in this country

11 Aventis does not waive and expressly reserves all rights to the confidentiality of data and information contained in all submissions to the FDA relating to Arava®, including, but not limited to, the NDA and the Investigational New Drug application for Arava®. 5 U.S.C. §552(b)(4), 21 C.F R §§312.130, 314.430, 20.61.

I. RHEUMATOID ARTHRITIS IS A SEVERE, DISABLING DISEASE; ALL AVAILABLE THERAPIES HAVE LIMITATIONS

HRG argues that: (i) methotrexate is the "gold standard" for the treatment of RA; (ii) methotrexate is more efficacious and safer than Arava®: and (iii) methotrexate and the other available therapies, as well as surgery, exercise, and rest, are adequate substitutes for Arava®.[12] These arguments minimize or ignore the severity of RA, disregard the risks of secondary failures and adverse events associated with methotrexate and other available therapies (including the recently approved biologic agents), and understate the benefits offered by Arava®.

A. THE SEVERITY AND PREVALENCE OF RHEUMATOID ARTHRITIS

Rheumatoid arthritis is a severe, chronic, debilitating disease where the body's immune system loses its normal regulatory mechanisms and attacks the healthy tissue lining the joints. This leads to inflammation of the joints and destruction of the adjoining soft tissues and bone, resulting in pain and loss of physical function. Joint damage can occur in the first year of the disease process, and the probability of developing erosions or other joint damage within the first 2 years is over 70 percent.[13] RA is progressive, often resulting in joint deformity and physical disability, and is associated with an increased risk of premature mortality. Because RA is a systemic disease, it causes fatigue and malaise and may also damage other organs, such as the heart, lungs, spleen and skin.

Approximately 2 million persons in the United States have RA, 70 percent of whom are women. Although the onset of disease frequently occurs in the 20s and 30s, with incidence and prevalence increasing with age, RA affects all age and ethnic groups in all parts of the world. The exact cause of RA is not known, and there is no known cure.[14]

---

12 The Petition suggests that methotrexate, the most widely used DMARD on the market (used for the past 25 years and approved for RA in 1986), is the drug of choice because of physician familiarity and fewer associated adverse events than other DMARDs. However, reporting of adverse events has been shown to significantly decrease after the first two years that a drug is on the market, because, among other things, physicians tend to report adverse events less frequently once they become familiar with use of individual medications (often referred to as the "Weber" effect). Tsong, Y Comparing reporting rates of adverse events between drugs with adjustment for year of marketing and secular trends in total reporting *J of Biopharm Stat*, 1995: 5(1): 95-114. It is impossible to directly compare a 25-year old drug with.a drug approved only 4 years ago when using spontaneous adverse event reporting as an index of safety.

13 ACR Guidelines Update 2002; Fuchs HA, et al. Evidence of significant radiographic damage in rheumatoid arthritis within the first 2 years of disease. *J Rheumatol* 1989;16(5).585-591; Brook A, et al. Radiographic changes in early rheumatoid disease *Ann Rheum Dis* 1977,36 71-73, Mottonen TT Prediction of erosiveness and rate of development of nbew erosions in early rheumatoid arthritis. *Ann Rheum Dis* 1988;47.648-653, Plant MJ, et al. Measurement and prediction of radiological progression in early rheumatoid arthritis. *J Rheumatol* 1994,21(10):1808-1813. An additional study found radiographic damage in 70% of patients within 3 years. Van der Heijde DMFM, et al Biannual radiographic assessments of hands and feet in a three-year prospective followup of patients with early rheumatoid arthritis. *Arthritis Rheum* 1992;35(1)·26-34.

14 Hochberg MC. Adult and juvenile rheumatoid arthritis: current epidemiologic concepts. *Epidemiol Rev* 1981; 3:27-44; Borigini MJ, et al. "Rheumatoid Arthritis In, Treatment of the Rheumatic Diseases: Companion to the Textbook of Rheumatology; Weisman MH and Weinblatt ME, eds. WB Saunders Co., Phila. c1995; Allaire S, et al. The costs of rheumatoid arthritis; *PharmacoEconomics* 1994, 6;(6):513-522;

1. Rheumatoid Arthritis Is Associated With Significant Adverse Health Consequences

Rheumatoid arthritis is associated with significant morbidity and premature death.[15] In a 25-year prospective study, median life expectancy of RA patients was shortened by 4 to 7 years in males and 3 to 10 years in females.[16] In RA patients with severely impaired physical function, 5-year survival was 50 percent or less, a prognosis no less severe than that of patients with Stage IV Hodgkin's lymphoma or 3-vessel coronary artery disease.[17]

RA may result in premature death due to complications of the disease in the joints or extra-articular (non-joint) manifestations of the disease, as discussed below:

- RA can lead to an unstable cervical spine and paralysis or death. Damaged joints can become infected, leading to potential infection in the bloodstream (i.e., septicemia or sepsis), which can be fatal.[18]

- Extra-articular manifestations of RA may include cardiac disease caused by rheumatoid inflammation of the heart lining (pericarditis), muscle (myocarditis), or valves (endocarditis); pulmonary disease (rheumatoid lung); vasculitis; amyloidosis, which can affect the kidneys; and Felty's syndrome, which can result in life-threatening infection. In addition, RA carries an increased risk of lymphomas and serious infections, such as pneumonia or sepsis, due to suppression of normal immune responses.[19]

- Premature coronary artery disease associated with RA is believed to be related to the B-cell, macrophage, and T-cell effects of this autoimmune disease.[20]

Goronzy J and Weyand C,. Rheumatoid Arthritis, Epidemiology, Pathology, and Pathogenesis. In. Primer on the Rheumatic Diseases. Klippel JH, ed. edition 12, Arthritis Foundation, Atlanta, 2001.

15 *See* footnote14; Wolfe F The burden of rheumatoid arthritis. *Am J Manag Care.* 1999; 5:S852-S859, Mikuls, TR, Saag, KG Comorbidity in rheumatoid arthritis. *Rheum Dis Clin North Am* 2001,27.283-303; Pincus T, et al. Premature mortality in patients with rheumatoid arthritis. evolving concepts. *Arthritis Rheum* 2001,44(6):1234-1236, Mitchell DM, et al. Survival, prognosis, and causes of death in rheumatoid arthritis. *Arthritis Rheum* 1986;29(6).706-714; Pincus T, et al. Taking mortality in rheumatoid arthritis seriously—predictive markers, socioeconomic status and comorbidity. *J Rheumatol* 1986,13(5).841-845; Scott DL, et al Long-term outcome of treating rheumatoid arthritis. results after 20 years *Lancet* 1987;1:1108-1111.

16 Vandenbroucke JP, et al. Survival and cause of death in rheumatid arthritis. A 25-year prospective followup. *J Rheum* 1984;11(2):158-161; Mitchell DM, et al. Survival, prognosis, and causes of death in rheumatoid arthritis. *Arthritis Rheum* 1986,29(6):706-714; Scott DL, et al. Long-term outcome of treating rheumatoid arthritis: results after 20 years *Lancet* 1987;1.1108-1111.

17 Pincus T, et al. Taking mortality in rheumatoid arthritis seriously—predictive markers, socioeconomic status and comorbidity. *J Rheumatol* 1986;13(5):841-845; Pincus T, et al. Prediction of long-term mortality in patients with rheumatoid arthritis according to simple questionnaire and joint count measures. *Ann Int Med* 1994;120(1).26-34

18 *See* footnote 16, Vandenbroucke, Goldenberg DL, Bacterial Arthritis In: Textbook of Rheumatology, 6th edition, Kelley WN, et al, eds , WB Saunders Co, Phila 2001; Harris ED. Clinical features of rheumatoid arthritis. In Textbook of Rheumatology, 6th edition, Kelley WN, et al, eds., WB Saunders Co, Phila 2001.

19 Borigini MJ, et al. "Rheumatoid Arthritis. In: Treatment of the Rheumatic Diseases Companion to the Textbook of Rheumatology; Weisman MH and Weinblatt ME, eds WB Saunders Co., Phila. c1995, Mitchell DM, et al Survival, prognosis, and causes of death in rheumatoid arthritis. *Arthritis Rheum* 1986;29(6).706-714; Vandenbroucke JP, et al Survival and cause of death in rheumatid arthritis: A 25-year prospective followup. *J Rheum* 1984;11(2)158-161, Wolfe F, et al. The mortality of rheumatoid arthritis. *Arthritis Rheum* 1994,37(4):481-494; Anderson R. Rheumatoid Arthritis, Clinical and Laboratory Features In Primer on the Rheumatic Diseases Klippel JH, ed. edition 12, Arthritis Foundation, Atlanta, 2001, Harris ED Clinical features of rheumatoid arthritis. In: Textbook of Rheumatology, 6th edition, Kelley WN, et al, eds., WB Saunders Co, Phila 2001.

20 Goodson N. Coronary artery disease and rheumatoid arthritis. *Current Opinion in Rheumatology* 2002; 14:115-120; Meyer O Atherosclerosis and connective tissue diseases *Joint Bone Spine* 2001;68:564-575; del Rincón I, et al. High Incidence of cardiovascular events

In addition, other extra-articular manifestations of RA may add to the chronic debility, such as rheumatoid eye disease (which can cause blindness); inflammation of tear glands and salivary glands (called Sjogren's syndrome) with ocular and oral complications; neuropathy; inflammatory nodules under the skin (called rheumatoid nodules); fatigue; fever; and anemia.[21]

Few RA patients have a short disease course with spontaneous and permanent remission. Most RA patients have progressive disease over the years, with periods of worsened disease activity (called flares).[22] The most advanced stages of RA are characterized by debilitation due to destruction of cartilage and bone and may include bony ankylosis (fusion) of the joint, joint deformity, and extensive muscle atrophy, with inability to perform even the most simple activities of daily living.[23]

2. Rheumatoid Arthritis Is Associated With Significant Economic And Personal Consequences

Impaired physical function associated with RA leads to decreased ability or inability to perform regular activities of daily living, work disability and reduced health-related quality of life. Work disability has been reported in 50 percent of RA patients within 10 years of onset of the disease.[24]

Rheumatoid arthritis accounts for over 250,000 hospitalizations and over 9 million physician visits yearly.[25] The costs to society have been estimated at up to $14 billion per year.[26] RA patients have 3 times the direct medical care costs, twice the hospitalization rate, and 10

---

in a rheumatoid arthritis cohort not explained by traditional cardiac risk factors. *Arthr Rheum* 2001;44(12).2737-2745; Manzi S, et al. Inflammation-mediated rheumatic diseases and atherosclerosis *Ann Rheum Dis* 2000; 59.321-325.

21 Anderson R Rheumatoid Arthritis, Clinical and Laboratory Features. In. Primer on the Rheumatic Diseases. Klippel JH, ed. edition 12, Arthritis Foundation, Atlanta, 2001; Harris ED. Clinical features of rheumatoid arthritis. In: Textbook of Rheumatology, 6th edition, Kelley WN, et al, eds., WB Saunders Co, Phila 2001

22 *Id.*, Harris; Matteson E. Rheumatoid Arthritis, Treatment. In: Primer on the Rheumatic Diseases. Klippel JH, ed. edition 12, Arthritis Foundation, Atlanta, 2001; Pope RM. Rheumatoid arthritis. pathogenesis and early recognition. *Am J Med* 1996;100 (Supp 2A)·3S-8S

23 Steinbrocker O, et al. Therapeutic criteria in rheumatoid arthritis *JAMA* 1949,140.659-662.

24 Felts W et al. The economic impact of the rheumatic diseases in the United States. *J Rheumatol* 1989; 16:867-884; Allaire SH, Prashker MJ, Meenan RF The costs of rheumatoid arthritis. Pharmaco-economics 1994, 6.513-22, Kobelt G, Eberhardt K, Jonsson L, Jonsson B. Economic consequences of the progression of rheumatoid arthritis in Sweden. *Arthritis Rheum* 1999;42:347-56; Pugner KM, Scott DI, Holmes JW, Hieke K The costs of rheumatoid arthritis: an international long-term view *Semin Arthritis Rheum* 2000;29.305-20; Yelin E, Callahan LF. The economic cost and social and psychological impact of musculoskeletal conditions. *Arthritis Rheum* 1995;38·1351-62; Yelin E, Wanke LA. An assessment of the annual and long-term direct costs of rheumatoid arthritis the impact of poor function and functional decline. *Arthritis Rheum* 1999; 42 1209-18, Yelin E The costs of rheumatoid arthritis: absolute, incremental, and marginal estimates. *J Rheumatol* 1996,23 suppl 44:47-51

25 *ACR Guidelines Update 2002;* Wong JB, Ramey DR, Singh G. Long-term morbidity, mortality, and economics of rheumatoid arthritis. *Arthritis Rheum* 2001;44.2746-49; van Jaarsveld CH, Jacobs JW, Schrijvers AJ, Heurkens AH, Haanen HC, Bijlsma JW. Direct cost of rheumatoid arthritis during the first six years. a cost-of-illness study. *Br J Rheumatol* 1998,37.837-47, Pincus T. The underestimated long-term medical and economic consequences of rheumatoid arthritis. *Drugs* 1995, 50 (suppl 1).1-14; Merkesdal S, Ruof J, Schoffski O, Bernitt K, Zeidler H, Mau W. Indirect medical costs in early rheumatoid arthritis. composition of and changes in indirect costs within the first three years of disease. *Arthritis Rheum* 2001; 44. 528-34.

26 Callahan L. The burden of rheumatoid arthritis. facts and figures. *J Rheumatol* 1998;25 (Suppl. 53):8-12.

times the work disability rate of an age- and sex-matched population.[27] Lost earnings for RA patients have been estimated to be $6.5 billion annually.[28]

B. THE STANDARD OF CARE AND THE LIMITED TREATMENT OPTIONS AVAILABLE FOR RHEUMATOID ARTHRITIS

1. Early And Aggressive Treatment With DMARDs Is The Standard Of Care

Because there is no known cure for rheumatoid arthritis, the ultimate goal in treating the disease is to induce a complete remission, which rarely occurs. More realistic goals of RA management are to control disease activity, alleviate pain, maintain physical function (especially to perform activities of daily living and work), maximize the patient's health related quality of life, and control or prevent joint damage. Because RA is a chronic progressive disease, proper management typically requires a lifelong coordinated effort involving medications, physical and occupational therapy, patient education, supportive services (when appropriate), and reconstructive surgery (when indicated). Periodic reassessment of disease activity, progression, and therapeutic efficacy, as well as vigilance to detect adverse effects, are essential and frequently require modification of the treatment regimen.[29]

RA treatment during the past 10 years has focused on early and aggressive use of DMARDs, which was primarily methotrexate until the past 4 years, when four new DMARDs and three new anti-inflammatory medications were approved. DMARDs interfere with the disease process and have the potential to modify the course of the disease. The ACR Guidelines recommend that DMARD therapy should be started within 3 months of diagnosis in the majority of patients with newly diagnosed RA. If repetitive flares occur, ongoing disease activity is present after 3 months of maximum therapy, or progressive joint damage is detected, then a switch to a different DMARD or addition of another DMARD is recommended.[30] Because not all patients have an adequate response to one DMARD alone, the use of combination DMARD therapy has come to play an important role in RA treatment.[31]

---

27 Felts W et al. The economic impact of the rheumatic diseases in the United States. *J Rheumatol* 1989; 16:867-884

28 *See* footnotes 26-27; ACR Guidelines Update 2002; Mitchell JM, et al. The importance of age, education, and comorbidity in the substantial earnings losses of individuals with symmetric polyarthritis. *Arthr Rheum* 1988;31(3):348-357.

29 ACR Guidelines Update 2002.

30 *Id.*

31 ACR Guidelines Update 2002, Kremer, JM Rational use of new and existing disease-modifying agents in rheumatoid arthritis. *Annals of Internal Medicine* 2001;134(8):695-703; Arava® (leflunomide) Prescribing Information; Pincus T, et al. Combination therapy with multiple disease-modifying antirheumatic drugs in rheumatoid arthritis: a preventive strategy *Annals Int Med* 1999; 131(10):768-774; Matteson

The DMARDs commonly used in RA treatment are Plaquenil (hydroxychloroquine), Azulfidine (sulfasalazine), methotrexate, Arava® (leflunomide), Enbrel® (etanercept), and Remicade® (infliximab).[32] In the United States, Plaquenil is often used as initial treatment in patients with early, mild disease.[33] Methotrexate is the most widely used DMARD; it is frequently added to Plaquenil and is the background DMARD for most DMARD combinations.[34] Remicade® is used in combination with methotrexate, after an inadequate response to methotrexate alone. Arava® and Enbrel® are used alone as alternatives to methotrexate or in combination with methotrexate after inadequate response to methotrexate alone.[35]

2. The Efficacy And Safety Limitations Of The Available DMARDS

HRG calls for the withdrawal of Arava® from the market given the availability of other therapies to treat RA. However, HRG does not objectively evaluate these alternatives relative to current medical knowledge and clinical practice. Some of the proposed alternatives are not viable for many patients. For example, HRG asserts that "rest and nutrition for acute attacks," "exercise," and "physiotherapy" are effective alternatives to Arava®. *See* Petition at 17. They are not alternatives to DMARDs, but provide only adjunct therapy. While exercise and rest are important additions to overall coordinated RA management and may help alleviate some symptoms, they are not the standard of care for treating active RA, because, above all, they do not prevent or slow progression of this disease. Surgery is also listed in the Petition as an alternative to Arava®. However, while surgery has a place as a reconstructive measure, it cannot be used to control disease activity. Even when indicated, it has its own attendant risks and is not an option for many arthritis patients due to their age, medical condition, or confounding factors associated with their disease.

E, Anderson R. Rheumatoid Arthritis In: Primer on the Rheumatic Diseases Klippel JH, ed. edition 12, Arthritis Foundation, Atlanta, 2001. DMARDs are often given concomitantly with other drugs used for symptomatic relief, typically· (1) nonsteroidal anti-inflammatory drugs (NSAIDs), including the new selective cyclooxygenase-2 (COX-2) inhibitors, and (2) low-dose corticosteroids (glucocorticoids), such as prednisone, which are potent anti-inflammatory drugs *Id*

32 ACR Guidelines Update 2002. Although highlighted by HRG, older DMARDs, such as gold preparations, penicillamine, cyclosporine, azathioprine, and cyclophosphamide, have only very limited use in current practice, particularly in the United States. *See* ACR Guidelines Update 2002. Sulfasalazine is widely used in Europe, but less so in the United States. Although Plaquenil (hydroxychloroquine) is categorized as a DMARD, there is no objective evidence that it modifies the course of disease progression. Plaquenil is often used as initial therapy in patients with mild RA

33 ACR Guidelines Update 2002; Matteson E, Anderson R. Rheumatoid Arthritis. In: Primer on the Rheumatic Diseases. Klippel JH, ed. edition 12, Arthritis Foundation, Atlanta, 2001.

34 ACR Guidelines Update 2002.

35 *Id.* Approximately 26 percent of Arava® use is in combination with methotrexate Scott-Levin's Physician Drug and Diagnosis Audit, 2001.

HRG's analysis of alternative medications is incomplete and out of date. Although most of the other medications mentioned have an important place in the treatment of RA, it is grossly misleading for HRG to suggest that their safety and efficacy profiles are superior to Arava®. Each has a different mechanism of action than Arava®, not all are equally beneficial in any individual patient, and all have been associated[36] with serious and sometimes fatal adverse events, many of which were not identified in the clinical trials leading to approval of the respective drugs.

Likewise, many RA patients have pre-existing co-morbid conditions that contraindicate the use of certain prescription medications.[37] Moreover, it is typical of the disease that patients become refractory to a particular medication, or that efficacy decreases over time. Although monotherapy with methotrexate is usually effective for some period of time, it is common for patients to eventually "fail" this therapy and require addition or substitution of treatments – thus, the recent reports of combination therapy in RA.[38] This well-known phenomenon is a fundamental reason underscoring the still unmet need for more RA medications with differing mechanisms of action.

In addition, each of the medications listed by HRG as alternative therapies has been associated with rare but serious adverse events that can be life-threatening or fatal.

- Methotrexate is associated with sometimes fatal pulmonary events (interstitial pneumonitis),[39] hepatic events (including cirrhosis), hematologic events (including pancytopenia, agranulocytosis and aplastic anemia), serious infections (including opportunistic infections), hemorrhagic enteritis, reversible renal

---

36 As used in this Response, the term "associated" refers to a temporal relationship between the medications and the events, not necessarily to a determination of causation. Robinson WH et al: Review: Demyelinating and neurological events reported in association with TNF antagonism: *Arth Rheum* 2001; 44.1977-83; ACR Hotline: FDA Advisory Committee reviews safety of TNF inhibitors, ACR 9/24/01, Keane J et al: Tuberculosis associated with infliximab, a TNFa neutralizing agent *N Engl J Med.* 2001, 345. 1098-104; Mohan N et al: Demyelination occurring during anti TNFa therapy for inflammatory arthritides *Arth Rheum* 2001; 44· 2862-9; Shakoor N et al: Drug induced systemic lupus erythematosus associated with etanercept therapy. *Lancet* 2002; 359:579-80.

37 *E.g.*, the new biologic Remicade® (infliximab) has been contraindicated in patients with congestive heart failure due to observed further deterioration in these patients during Phase II clinical trials.

38 Kremer, JM and Lee, JK, A long-term prospective study of the use of methotrexate in rheumatoid arthritis: Update after a mean of 53 months, *Arth and Rheum*, 1988, 31.577-584. O'Dell J, Haire C, Erikson N et al: Treatment of rheumatoid arthritis with methotrexate alone, sulfasalazine and hydroxychloroquine, or a combination of all three. *N Engl J Med* 1996, 334: 1287-91 Tugwell P, Pincus T, Yocum D et al: Combination therapy with cyclosporin and methotrexate in severe RA. *N Engl J Med* 1995; 333:137-141. Weinblatt ME, Kremer JM, Bankhurst AD, Bulpitt KJ et al: A trial of etanercept, a recombinant tumor necrosis factor receptor-Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. *N Engl J Med.* 1999; 340:253-9 Lipsky PE, van der Heijde, DM, St Clair EW, Furst, DE, Breedveld FC, Kalden JR, et al Infliximab and Methotrexate in the Treatment of Rheumatoid Arthritis *N Engl J Med.* 2000;343:1594-1602. Cohen S., Hurd E., Cush J , Schiff M., Weinblatt ME., Moreland LW et al. Treatment of rheumatoid arthritis with anakinra, a recombinant human interleukin-1 receptor antagonist (IL-1ra), in combination with methotrexate *Arth Rheum* 2001; 46 614-24 Weinblatt ME, Kremer JM, Coblyn JS, Maier AM, Helfgott SM, Morrell M, Byrne VM, Kaymakcian MV, Strand V: Efficacy, Safety and Pharmacokinetics of the combination of methotrexate and leflunomide in patients with active rheumatoid arthritis *Arth Rheum* 1999; 42:1322-8. KremerJM, Genovese MC, Cannon GW et al. Concomitant leflunomide therapy in patients with active rheumatoid arthritis despite stable doses of methotrexate: A randomized comparison of efficacy, safety, and tolerability compared to methotrexate alone *Ann Int Med*, (accepted for publication).

39 Zisman, DA, et al., Drug-induced pneumonitis. the role of methotrexate, *Sarc Vasc and Diff Lung Dis 2001;* 18(3): 243-252; Cannon GW Cerveny KC, Finck BK Enbrel ERA Investigators Group, Simpsn KM, Leflunomide Investigators Group, Strand V; Incidence and Risk Factors for Methotrexate-induced Pulmonary Disease during Treatment of Rheumatoid Arthritis. *Arthritis Rheum* 44: S341.

failureand severe skin reactions (including Stevens-Johnson syndrome and toxic epidermal necrolysis).[40]

- Azulfidine (sulfasalazine) is associated with sometimes fatal hematologic events (agranulocytosis and aplastic anemia), renal and hepatic damage, hypersensitivity reactions, irreversible neuromuscular and central nervous system changes, and fibrosing alveolitis. Severe skin hypersensitivity reactions have included Stevens-Johnson syndrome and toxic epidermal necrolysis. Hepatic events have included hepatitis, jaundice, cholestatic jaundice, fulminant hepatitis, hepatic necrosis, hepatic failure, and cirrhosis. Hemolytic anemia can occur in patients with underlying glucose-6-phosphate deficiency.[41]

- Enbrel® (etanercept) is associated with aplastic anemia, demyelinating neurologic diseases, tuberculosis, other opportunistic infections, and fatal cases of sepsis. Enbrel® is also associated with various opportunistic infections, including nasal and bronchial infections, *staphylococcus auereus* infection, and *E. coli* urinary tract infections.[42]

- Remicade® (infliximab) is associated with active tuberculosis, as well as other opportunistic infections, and deaths due to sepsis and tuberculosis.[43]

- Plaquenil® (hydroxychloroquine) is associated with rare irreversible retinal damage which can lead to visual loss. With overdose or with lower doses in hypersensitive patients, sudden respiratory and cardiac arrest has occurred. It is also associated with neuromuscular reactions, serious skin reactions such as Stevens-Johnson syndrome and exfoliative dermatitis, and hematologic events including aplastic anemia, granulocytosis, leukopenia, thrombocytopenia, and hemolysis in individuals with glucose-6-phosphate deficiency.[44]

- Injectable gold is associated with anaphylactic shock, hematologic events (aplastic anemia, hypoplastic anemia, agranulocytosis, pancytopenia, leukopenia, thrombocytopenia, hemorrhagic diathesis), hepatic events (cholestasis, toxic hepatitis, jaundice) interstitial pneumonitis or fibrosis, renal disease, and severe skin reactions such as exfoliative dermatitis. [45]

- Penicillamine is associated with sometimes fatal aplastic anemia, agranulocytosis, thrombocytopenia, Goodpasture's syndrome (a severe pulmonary-renal disease),

---

40 Methotrexate prescribing information, revised August 28, 2001

42 Azulfadine® EN-tabs (sulfasalazine delayed-release tablets) prescribing information.

42 Enbrel® prescribing information; Ferraccioli G et al , Anticardiolipin antibodies in rheumatoid patients treated with etanercept or conventional combination therapy: direct and indirect evidence for a possible association with infections, *Annals of Rheumatic Disease; 2002* 61(4): 358-61; Mohan N et al: Demyelination occurring during anti TNFa therapy for inflammatory arthritides Arth Rheum 2001; 44: 2862-9. Shakoor N et al: Drug induces systemic lupus erythematosus associated with etanercept therapy. Lancet 2002; 359:579-80. Robinson WH et al: Review: Demyelinating and neurological events reported in association with TNFa antagonism: Arth Rheum 2001; 44:1977-83. ACR Hotline: FDA Advisory Committee reviews safety of TNF inhibitors, ACR 9/24/01 www.fda.gov/medwatch/safety/1999/enbrel.html.

43 Remicade® prescribing information, Keane J, et al , Tuberculosis associated with infliximab, a tumor necrosis factor alpha-neutralizing agent, *N Engl J Med.; 2001* 55 (15). 1098-104; ACR Hotline FDA Advisory Committee reviews safety of TNF inhibitors, ACR 9/24/01. www.fda.gov/medwatch/safety/2000/remicade.html.

44 Plaquenil® (hydroxychloroquine) prescribing information

45 Solganol® (aurothioglucose) prescribing information.

and myasthenia gravis. Serious events also include renal disease, toxic hepatitis, drug-induced lupus erythematosus, neuropathy, and severe skin reactions including pemphegus vulgaris, exfoliative dermatitis, and toxic epidermal necrolysis.[46]

- Imuran (azathioprine) is associated with severe leukopenia and thrombocytopenia, serious infections including opportunistic infections, malignancies, hypersensitivity reactions, hepatic events, and interstitial pneumonitis.[47]

- Neoral (cyclosporine) is associated with dose-related renal toxicity (including structural kidney damage) and with hepatic events. Hypertension is common. Higher doses used in organ transplantation may increase the susceptibility to infection and neoplasia. Cyclosporine has known interactions with many drugs such as various antibiotics, anti-fungals, anti-neoplastics, anti-inflammatory drugs such as NSAIDs and methylprednisolone, anti-convulsants, and histamine-2 receptor blockers.[48]

- NSAIDs, widely used in the symptomatic treatment of many acute and chronic inflammatory and painful conditions, are associated with sometimes fatal complications of peptic ulcer disease (especially gastrointestinal hemorrage), fatal anaphylactoid reactions, and hepatic events including hepatic necrosis, jaundice, and fulminant hepatitis. In addition, NSAIDs are associated with renal damage, aseptic meningitis with fever and coma, hematologic events (including leukopenia, thrombocytopenia, aplastic anemia, agranulocytosis, and hemolytic anemia), and severe skin reactions such as Stevens-Johnson syndrome.[49]

- Glucocorticoids ("corticosteroids", most often prednisone) are potent anti-inflammatory drugs used widely in the treatment of RA and other diseases. They are associated with many adverse events, especially with long-term use, even at the lower doses (≤10 mg/day) usually used in RA, including increased susceptibility to and seriousness of infections (including opportunistic infections), cardiac ventricular wall rupture after recent myocardial infarction, and acute adrenal insufficiency in physiologic stress situations or with abrupt cessation of the glucocorticoid.[50] Other potentially serious adverse events include diabetes, hypertension, atherosclerosis, osteoporosis leading to fractures, a type of joint damage called avascular necrosis, glaucoma, and impaired wound healing.[51]

In short, use of every therapy currently indicated for the treatment of active RA has both recognized benefits and risks. Aventis does not refer to other RA medications for comparison

---

46 Cuprimine® (penicillamine) prescribing information.

47 Imuran® (azathioprine) prescribing information.

48 Neoral® (cyclosporine microemulsion) prescribing information

49 NSAID class labeling, e.g., Voltaren® (diclofenac) prescribing information; Tolman KG, Hepatoxicity of Non-Narcotic Analgesics, *Am J of Med*; 1998 105: 13S-19S.

50 Glucocorticoid class labeling, e.g Deltasone® (prednisone) prescribing information.

purposes or to suggest that they are unsafe, but rather to offer perspective. These products, including Arava®, are indicated for treatment of a chronic disease with devastating complications, which can be life threatening. None of these therapies is risk free; however, the benefits of each outweigh their associated risks. Because RA is heterogeneous and affects each patient differently, each treatment must be individually selected for optimal use in each patient at their particular stage of the disease process. The balance of this Response will evaluate the efficacy and safety of use of Arava® and confirm its positive benefit-risk profile in the treatment of RA.

II. ARAVA® IS A NOVEL THERAPY IN THE TREATMENT OF RHEUMATOID ARTHRITIS

Arava® is an isoxazole immunomodulatory agent with a unique mechanism of action. It inhibits *de novo* pyrimidine syntheses by reversibly blocking the enzyme dihydroorotate dehydrogenase (DHODH), resulting in antiproliferative effects on activated autoimmune lymphocytes important in the pathogenesis of RA.[52]

The NDA for Arava® was submitted to the FDA on March 10, 1998. Because Arava® was judged to offer a new therapeutic alternative in a debilitating, potentially life threatening disease, it was assigned priority review by the FDA in April 1998, and given a 1P designation, which is reserved for drugs that, if approved, would represent a significant improvement compared to marketed products in the treatment, diagnosis, or prevention of a specific disease.[53] The FDA Arthritis Advisory Committee unanimously recommended approval on August 7, 1998,[54] and the NDA was approved on September 10, 1998. Arava® was the first DMARD indicated to retard structural damage as evidenced by x-ray erosions and joint space narrowing. Since 1998, Arava® has been used by over approximately 500,000 patients worldwide.

---

51 ACR Ad Hoc Committee on Clinical Guidelines. Guidelines for monitoring drug therapy in Rheumatoid Arthritis. *Arthritis Rheum* 1996,39(5):723-731.

52 Fox RI, et al. Short analytical review. mechanism of action for leflunomide in rheumatoid arthritis. *Clin Immunol* 1999; 93(3) 198-208; Fox RI. Mechanism of action of leflunomide in rheumatoid Athritis. *J Rheumatol* 1998; 25 Supp 53:20-26, Kremer JM. Methotrexate and leflunomide: Biochemical basis for combination therapy in the treatment of rheumatoid arthritis. *Seminars in Arthritis and Rheumatism* 1999;29(1):14-26, Breedveld FC, et al. Leflunomide: mode of action in the treatment of rheumatoid arthritis. *Ann Rheum Dis* 2000; 59:841-849; Laan R, et al. Leflunomide and methotrexate. *Current Opinion in Rheumatology* 2001; 13. 159-163

53 See FDA/CDER Manual of Policies & Procedures, Priority Review Policy, MAPP 6020 3.

54 FDA Talk Paper, T98-54, September 11, 1998.

A. THE PETITION MISREPRESENTS THE EFFICACY OF ARAVA® IN CLINICAL TRIALS

A constant (but unsubstantiated) theme in the Petition is that Arava® is inferior to methotrexate and, at a minimum, is not equally effective compared with methotrexate. This claim is erroneous in both premise and fact. First, the length of time on any DMARD, including methotrexate, is ultimately limited by intolerance and/or loss of effectiveness.[55] Therefore, RA patients typically need to take many different DMARDs during the course of their disease.[56] It is, therefore, essential to have multiple alternatives for monotherapy, as well as options for various DMARD combinations.[57] Moreover, no recent DMARD approval (including Arava® and the biologic agents) for the treatment of active RA has been based on the demonstration of superior efficacy compared with methotrexate.

Second, as discussed below, the controlled clinical trial data demonstrate that, overall, Arava® and methotrexate have equivalent efficacy without consistent or meaningful clinical differences across studies. Third, the clinical trial data were extensively reviewed by the FDA and its Arthritis Advisory Committee prior to approval. HRG not only misrepresents the efficacy and safety data of Arava®, but it does not provide any new information to suggest that either the FDA or the Advisory Committee made a decision based on incomplete or incorrect information.

B. CLINICAL TRIALS ESTABLISHED THE EFFICACY OF ARAVA®

Arava® was studied in randomized, controlled clinical trials involving more than 2400 adult patients before it was first approved for use by a regulatory health authority. Three phase III controlled clinical trials (each of which was continued in blinded extension trials) established the efficacy of Arava® in reducing the signs and symptoms of RA, improving physical function, and retarding structural joint damage:

- **US301** was a randomized, double-blind, placebo-controlled study of 482 patients, with a primary endpoint at 12 months and continued double-blind treatment to 24 months. Leflunomide was compared with both placebo and methotrexate (plus folate). The ACR 20 Responder-at-Endpoint rates were statistically equivalent for leflunomide (41%) and

55 Wolfe F. The epidemiology of drug treatment failure in rheumatoid arthritis, *Baillier's Clinical Rheumatology* 1995; 9(4):619-632.
56 ACR Guidelines Update 2002.
57 Contrary to the Petition, the value of any DMARD, including Arava®, in the treatment armamentarium is not based on demonstration of increased efficacy compared to methotrexate

methotrexate (35%), and both were statistically significantly superior to placebo (19%).[58] Leflunomide and methotrexate were statistically significantly better than placebo by ACR 20 rates, including tender and swollen joint counts, global assessments, pain, ESR, CRP,[59] and physical function and health related quality of life assessments. Onset of action was faster with leflunomide, and leflunomide resulted in greater improvement in HAQ Disability Index, *see infra* subsectionII.B.2., and equivalent slowing or inhibiting radiographically-assessed disease progression compared with methotrexate. The second year data showed maintenance of clinical and radiographic benefits at 24 months in both active treatment groups.[60]

- **MN301** was a randomized, double-blind, placebo-controlled, 6-month study of 358 patients, and the active comparator drug was sulfasalazine. The ACR 20 Responder-at-Endpoint rates were 49% for leflunomide, 45% for sulfasalazine, and 29% for placebo. Leflunomide and sulfasalazine were statistically equivalent and both were statistically significantly superior to placebo by ACR 20 rates and all components, HAQ Disability Index, and slowing or inhibiting radiographic disease progression.[61]

- **MN302** was a 999-patient randomized, double-blind study comparing leflunomide to methotrexate at 12 months. This study was not placebo-controlled, and concomitant folate administration was not required (only 10% of methotrexate patients received folate).[62] The ACR 20 Responder-at-Endpoint rate was 43% with leflunomide and 57% with methotrexate, which showed statistical non-equivalence; the differences in the components of the response criteria were small and not considered clinically meaningful. In addition, the two treatments were statistically equivalent for slowing or inhibiting disease progression by x-ray and in

---

58 This stringent primary analysis is described in more detail, *infra* at subsection II.B.1., but, in short, measures the percentage of patients that had an ACR 20 Response and completed the trial Another analysis, the Last Observation Carried Forward ("LOCF"), measures the percent of patients that had an ACR 20 Response whenever they discontinued in the study, or at the end if they remained in the study to completion. The ACR 20 Response rates in the LOCF analysis in US301 were similar or leflunomide (52%) and methotrexate (46%), and both were superior to placebo (26%).

59 ESR is erythrocyte sedimentation rate and CRP is C-reactive protein. Higher levels of either of these blood tests reflect degree of inflammation.

60 Strand V, et al. Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate. *Archives Int Med* 1999;159:2542-2550, Sharp JT, et al. Treatment with leflunomide shows radiographic progression of rheumatoid arthritis. *Arthritis Rheum* 2000,43.495-505; Cohen S, et al Two-year, blinded, randomized, controlled trial of treatment of active rheumatoid arthritis with leflunomide compared with methotrexate. *Arthritis Rheum* 2001;44(9):1984-1992.

61 Smolen JS, et al. Efficacy and safety of leflunomide compared with placebo and sulphasalazine in active rheumatoid arthritis: a double blind, randomized, multicentre trial. *Lancet* 1999,353.259-66, Arava® (leflunomide) prescribing information; Sharp JT, et al. Treatment with leflunomide shows radiographic progression of rheumatoid arthritis *Arthritis Rheum* 2000,43:495-505 The ACR 20 Response rates in the LOCF analysis were 55% for leflunomide, 57% for sulfasalazine, and 29% for placebo.

62 It is believed that folate decreases methotrexate toxicity, especially gastrointestinal symptoms and liver enzyme elevations (often called liver function tests or LFTs). ACR Ad Hoc Committee on clinical guidelines. Guidelines for monitoring drug therapy in Rheumatoid Arthritis. *Arthritis Rheum* 1996;39(5) 723-731, Van Ede AE, et al. Effect of folic or folinic acid supplementation on the toxicity and efficacy of methotrexate in rheumatoid arthritis. *Arthritis Rheum* 2001; 44(7): 1515-1524. . Furst DE, Cohen S, Emery P et al: Does Folic Acid decrease the efficacy as well as the toxicity of methotrexate in RA. *Arth Rhuem* 2001; 45 S373. Folate supplementation was mandated in US301, whereas only 10 percent of methotrexate patients in MN302 received folate supplementation.

an AUC analysis of ACR Response. *See infra* subsection II.B.1. Leflunomide had a more rapid onset of response. [63]

- **MN303** was a double-blind, 6-month extension of MN301. ACR 20 responses, x-ray benefit and improvements in physical function were maintained over 12 months in both leflunomide and sulfasalazine patients, and the two treatments were statistically equivalent in all clinical parameters studied.

- **MN305** was a double-blind extension of MNN301/303 for a second year. ACR 20 responses, x-ray benefits and improvements in physical function were maintained in year 2 in patients continuing leflunomide treatment. At month 24, leflunomide-treated patients had statistically significantly better ACR 20 Response rates, investigator and patient global assessments, HAQ Disability Index scores and x-ray benefit than sulfasalazine ; other efficacy parameters were similar in both treatment groups.[64]

- **MN304** was a double-blind extension of MN302 for a second year. ACR 20 responses, x-ray benefits and improvements in physical function were maintained in patients continuing a second year of leflunomide treatment. After 2 years of treatment, leflunomide and methotrexate had equivalent clinical efficacy by ACR Responses and HAQ Disability Index.[65]

These trials provide clear evidence of the important benefits provided by Arava®. Although this data were reported to the FDA in detail in the NDA and published in peer-reviewed journals, HRG failed to reference much of it – especially when the data were inconsistent with its position. The following discussion provides additional evidence of the proven efficacy of Arava®.

1. **Reduction In Signs And Symptoms**

The FDA requires that clinical efficacy for the treatment of RA be measured using a defined composite index of multiple signs and symptoms, such as determining the proportion of patients who meet the American College of Rheumatology ("ACR") criteria defining a clinical

---

63 Emery P, et al. A comparison of the efficacy and safety of leflunomide and methotrexate for the treatment of rheumatoid arthritis. *Rheumatology* 2000;39:655-665; *Arava*® (leflunomide) prescribing information, Sharp JT, et al. Treatment with leflunomide slows radiographic progression of rheumatoid arthritis. *Arthritis Rheum* 2000;43(3):495-505. The ACR 20 response rate (for LOCF) was 51% with leflunomide and 65% with methotrexate.

64 Kalden JR, et al. Improved functional ability in patients with rheumatoid arthritis—long-term treatment with leflunomide versus sulfasalazine. *J Rheum* 2001;28(9): 1983-91; Scott DL, et al. Treatment of active rheumatoid arthritis with leflunomide: two year follow up a double blind, placebo controlled trial versus sulfasalazine *Ann Rheum Dis.* 2001;60.913-923.

65 Emery P, et al. A comparison of the efficacy and safety of leflunomide and methotrexate for the treatment of rheumatoid arthritis. *Rheumatology* 2000;39.655-665

response, known as an ACR 20 Responder.[66] An ACR 20 Responder must have at least 20 percent improvement demonstrated in 5 of 7 core set measures of disease activity, including both tender and swollen joint counts.[67] Using these criteria, Aventis applied a stringent primary analysis -- the ACR 20 Responder-at-Endpoint rate – to the phase III clinical trial data.[68] In both placebo-controlled trials, Arava® monotherapy was statistically significantly superior to placebo in reducing the signs and symptoms of RA after 6 months in MN301 (ACR Responder-at-Endpoint: Arava®-49% vs. placebo-29%),[69] and after 12 months in US301 (41% vs. 19%), and statistically equivalent to the active comparator agents (methotrexate and sulfasalazine).[70]

HRG cites only the 12-month efficacy data from the MN302 study, where a difference in ACR 20 Responder-at-Endpoint rate was observed in favor of methotrexate (57%) over leflunomide (43%), although, as previously noted, the differences in the components were small and not meaningfully different from a clinical standpoint. However, HRG fails to reference the 12-month results of US301, in which there was no statistically significant difference in the ACR 20 Responder-at-Endpoint rate between Arava® (41%) and methotrexate (35%).[71] Indeed, the efficacy of Arava® was consistent across all trials, whereas the efficacy of methotrexate varied substantially between trials.[72] In addition to the efficacy demonstrated by ACR Response rates,

---

66 FDA Guidance to Industry: Clinical Development Programs for Drugs, Devices, and Biological Products for the Treatment if Rheumatoid Arthritis. *Clin* 1999 8.1-56, Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid Arthritis Clinical Trials Arthritis Rheum. 1993,36.729-740.

67 The core set measures used to determine whether a patient is a responder are tender and swollen joint counts, physician and patient assessments of disease activity, laboratory measures of disease activity (sedimentation rate or C-reactive protein), pain, and patient-reported assessment of physical function using a validated physical function instrument such as the Health Assessment Questionnaire. Felson DT, et al American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995, 38(6) 727-735. An ACR 20 Responder may also meet criteria for higher thresholds of response, an ACR 50 or ACR 70 Responder is defined in a manner analogous to the ACR 20 Responder but with improvements of at least 50% or at least 70%, respectively.

68 The primary efficacy analysis for overall clinical response in the Arava trials was the ACR 20 Responder-at -Endpoint analysis, a stringent analysis in which an ACR 20 Responder-at-Endpoint is a patient who both (1) completed the study and (2) was an ACR 20 responder at the study endpoint. Additionally, dropouts for any cause were considered non-Responders, even if they had an ACR 20 Response at the time they left the study. Each trial was extended to a total of 2 years and demonstrated that the benefit at 1 year was maintained in year-2.

69 Smolen JS, et al. Efficacy and safety of leflunomide compared with placebo and sulphasalazine in active rheumatoid arthritis: a double blind, randomized, multicentre trial. *Lancet* 1999;353:259-66.

70 Strand V, et al. Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate *Archives Int Med* 1999;159.2542-2550. MN302 was not a placebo-controlled trial.

71 Arava® (leflunomide) prescribing information. There was also no statistically significant difference in ACR 20 Responder-at-Endpoint rate between leflunomide (49%) and sulfasalazine (45%) after 6 months in MN301. *Id.*

72 In the ACR 20 Responder-at-Endpoint analysis across all trials, 41-49% of the Arava®-treated patients completed the 6- or 12-month trial with at least a 20% response at the end of the trial. In the LOCF analysis, using the last study visit for patients who discontinued early, more than half of the Arava®-treated patients (51-55%) had at least a 20% response at their last study visit, one third (31-34%) had at least a 50% response, and 10-20% had at least a 70% response

Unlike with Arava®, methotrexate efficacy varied considerably between trials In US301 and MN302, rates ranged from 35% to 57% for the ACR 20 Responder-at-Endpoint rates, and in the LOCF analysis, ranged from 46% to 65% for ACR 20, 23% to 44% for ACR 50 and 9-16% for ACR 70 Responder rates. The reasons for this variability in methotrexate performance between US301 and MN302 are not clear, but it may have been influenced by differences in patient populations, absence of a placebo arm, and the fact that folate supplementation was used in only 10 percent of methotrexate of patients in MN302. Emery P, et al. A comparison of the efficacy and safety of leflunomide and methotrexate for the treatment of rheumatoid arthritis. *Rheumatology* 2000;39 655-665, Strand V, et al Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate. *Archives Int Med* 1999, 59 2542-2550. *See also supra,* fn. 62.

treatment with Arava® improved all of the individual components of disease activity consistently across the three trials.[73] It should be noted that the use of methotrexate with folate in US 301 most closely mirrors how that drug is prescribed and used in the United States.

Moreover, when the ACR 20 Response was analyzed over time, Arava® and methotrexate were not statistically different in either US301 or MN302. Whereas the ACR 20 analysis measures a response at one point in time, the Area Under the Curve (AUC) analysis measures the number of weeks a patient is an ACR 20 Responder, which provides important detail regarding the onset and time course of patient response. AUC analyses showed statistical equivalence between Arava® and methotrexate in US301 and MN302 and equivalence between Arava® and sulfasalazine in MN301.[74]

Analyses of response over time also demonstrated that the treatment effect of Arava® was rapid and sustained. Response was evident by 1 month, with further increases, which stabilized by 3-6 months and continued throughout the course of treatment.[75] In patients with pain and inflammation, the time to onset of effect is an important consideration. Initial response and sustained response occurred earlier with Arava® compared with methotrexate in both studies.[76]

2. Improvement In Physical Function

Impairment in physical function may make it difficult to perform activities of daily living, resulting in work disability for many patients, and reducing health-related quality of life.[77] Maintaining physical function for activities of daily living and work, as well as health related quality of life, are important goals in RA management.[78]

---

73 Arava® (leflunomide) prescribing information

74 Strand V, et al. Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate. *Archives Int Med* 1999,159:2542-2550; Arava NDA 20-905

75 Arava® (leflunomide) prescribing information.

76 Strand V, et al Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate. *Archives Int Med* 1999;159·2542-2550; Emery P, et al A comparison of the efficacy and safety of leflunomide and methotrexate for the treatment of rheumatoid arthritis *Rheumatology* 2000;39.655-665. ACR 20 response was apparent at one month in 38% of patients treated with Arava compared to 24% with methotrexate in US301, 24% with Arava compared to 18% with methotrexate in MN302, and 31% with Arava compared to 19% with sulfasalazine in MN301. Arava® NDA 20-905.

77 Wolfe F, et al The clinical value of the Stanford Health Assessment Questionnaire Functional Disability Index in patients with rheumatoid arthritis *J Rheumatol* 1998;15(10).1480-1488, Strand V, 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(9):1870-8

78 ACR Guidelines Update 2002.

Physical function was assessed by the HAQ Disability Index -- a recognized, validated instrument used to assess rheumatic disease-specific impairment.[79] The HAQ was completed by all patients in all phase III clinical studies. The ACR Response criteria were calculated using mean HAQ (MN301 and MN302) or modified HAQ (US301), as well as patient global assessment and patient assessment of pain. HRG dismisses the HAQ analysis, concluding without discussion that it measures primarily subjective endpoints, *see* Petition at 15, and disregarding patient perception entirely. In fact, impairment of physical function has predictive value for work and overall disability, cost, joint replacement surgery, and premature mortality.[80] In the clinical trials, treatment with Arava® resulted in statistically significant improvement compared with placebo in the HAQ Disability Index, as well as all 8 HAQ subscale scores in both phase III placebo-controlled trials.[81] In all trials, improvement in HAQ Disability Index subscales in the leflunomide treatment groups was clinically meaningful and, in most of the subscales, exceeded or approached twice the minimal clinically important difference established in the literature at 6, 12 and 24 months.[82] These data show that Arava® did not merely maintain

---

79 The HAQ was developed to assess disease-specific physical function and degree of disability in patients suffering from RA. It consists of various questions relating to eight categories (dressing and grooming, rising, eating, walking, hygiene, reach, grip, and activities). Fries J, et al. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23(2).137-145; Ramey DR et al. The Health Assessment Questionnaire 1995—Status and Review In. Quality of Life and Pharmacoeconomics in Clinical Trials, second edition, Spilker B, ed. Lippencott-Raven Publ, PA, c1996.

80 Mitchell DM, et al. Survival, prognosis, and causes of death in rheumatoid arthritis. *Arthritis Rheum* 1986,29(6):706-714; Pincus T, et al. Prediction of long-term mortality in patients with rheumatoid arthritis according to simple questionnaire and joint count measures *Ann Int Med* 1994;120(1).26-34; Wolfe F, et al The long-term outcomes of rheumatoid arthritis. *Arthr Rheum* 1998; 41(6):1072-1082, Wolfe F, et al. Clinical and health status measures over time: prognosis and outcome assessment in rheumatoid arthritis *J Rheumatol* 1991,18(9):1290-1297, Wolfe F. The prognosis of rheumatoid arthritis assessment of disease activity and disease severity in the clinic. *Am J Med* 1997;103:12S-18S, Wolfe F, et al The clinical value of the Stanford Health Assessment Questionnaire Functional Disability Index in patients with rheumatoid arthritis. *J Rheumatol* 1988;15(10) 1480-1488; Fries JF, et al. Medical costs are strongly associated with disability levels in rheumatoid arthritis. *Arthritis Rheum* 1995;38(suppl.):S187; Singh G, et al. Long-term medical costs and outcomes are significantly associated with early changes in disability in rheumatoid arthritis. *Arthritis Rheum* 1996;39(suppl.):S318.

81 Strand V, 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(9) 1870-8, Tugwell P, et al Clinical improvement as reflected in measures of function and health-related quality of life following treatment with leflunomide compared with methotrexate in patients with rheumatoid arthritis. *Arthritis Rheum* 2000;43(3) 506-514; Kalden JR, et al Improved functional ability in patients with rheumatoid arthritis—long-term treatment with leflunomide versus sulfasalazine. *J Rheum* 2001,28(9): 1983-91

82 Wells G, et al. Important difference between patients with rheumatoid arthritis. the patient's perspective. *J Rheumatol* 1993;20:557-560. Kosinski M, Zhao SZ, Didhiya S, Osterhaus JT, Ware JE. Determining minimum clinically important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. *Arth Rheum* 2000;43 1478-87. Kujawski SC, Kosinski M, Martin R, Wanke LA, Buatti MC, Ware JE, et al. Determining meaningful improvement in SF-36 scale scores for treatment studies of early, active RA:.*Arth Rheum* 2000; 43.S140 Samsa G, Edelman D, Rothman M, et al. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. *Pharmacoeconomics* 1999; 15.141-155 Tugwell P, Wells G, Strand V, Bombardier C, Maetzel A, Crawford B, Dorner C, Thompson A: Clinical Improvement as Reflected in Measures of Function and health-related quality of life. Sensitivity and Relative Efficiency to Detect a Treatment Effect in a 12 month Placebo Controlled Trial Comparing Leflunomide with Methotrexate, *Arth Rheum* 2000; 43 506-14 Strand V, Cannon G, Cohen S, Ware J et al: Correlation of HAQ with SF-36· Comparison of Leflunomide to Methotrexate in patients with active RA. *Arth Rheum* 2001; 44·S187 Strand V, Bombardier C, Maetzel A, Scott D, Crawford B: Use of minimum clinically important differences [MCID] in evaluating patient responses to treatment of RA. *Arth Rheum* 2001; 44·S187. Zhao SZ, McMillen JI, Markensop JA, Dedhiya SD, Zhao WW, Osterhaus JT, Yu SS. Evaluation of the functional status aspects of health-related quality of life of patients with osteoarthritis. *Pharmacotherapy* 1999; 19.1269-1278 Ehrich EW, Bolognese JA, Kong S, Watson DJ, Zeng K, Seidenberg BC. Improvements in SF-36 mental health domains with treatment of OA result of decreased pain and disability or independent mechanism? *Arth Rheum* 1998, 41.S221 Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N: Minimum perceptible clinical improvement with the WOMAC and global assessments in patients with osteoarthritis. *J Rheumatol* 2000; 27·2635-41. Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their

the physical function present at baseline, but actually improved it to a statistically and clinically meaningful degree.

In addition to the HAQ Disability Index, two other instruments were used in US301 to further evaluate physical function and health related quality of life, neither of which is mentioned by HRG. One method -- the Problem Elicitation Technique (PET) questionnaire -- is based on the patient identifying those physical activities that he or she considers most important (i.e., activities that are most affected by their disease and that they would most want to see improved).[83] In this analysis, patients treated with Arava® showed statistically significantly greater improvement compared with both placebo and methotrexate treatment groups.[84]

The second additional method used to evaluate improvement in physical function in US301 was the SF-36 -- a widely used instrument to assess generic health-related quality of life. This was the first randomized clinical trial to demonstrate reduction in all domains of health related quality of life in RA patients compared to the general population (age and gender matched).[85] Arava® treatment resulted in statistically significant improvements compared to placebo in the Physical Component Summary score and in 5 of the 8 SF-36 domains (physical functioning, body pain, general health perception, vitality, and social functioning). Arava® also was associated with statistically significant improvement in the Physical Component Summary score and in 2 SF-36 domain scores (body pain and vitality) compared with methotrexate.[86] As with the HAQ, the PET and SF-36 instruments are recognized as important instruments to assess

---

implications for required sample sizes using WOMAC and SF-36 Quality of life measurement instruments in patients with osteoarthritis of the lower extremities *Arth Care Res* 2001. 45:384-391. Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD: Linking clinical relevance and statistal significance in evaluating intra-individual changes in HRQOL. Medical Care 1999, 37:469-78. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM based criterion for identifying meaningful intra-individual changes in HRQOL· *J Clin Epidemiol* 1999; 52:861-73. Kosinski M, Zhao SZ, Didhiya S, Osterhaus JT, Ware JE. Determining minimum clinically important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. *Arth Rheum* 2000;43:1478-87

83 Tugwell P, et al Methotrexate in rheumatoid arthritis: Impact on quality of life assessed by traditional standard-item and individualized patient preference health status questionnaires. *Arch Intern Med* 1990; 150:59-62-62.

84 Strand V, et al. Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate. *Archives Int Med* 1999,159·2542-2550; Strand V, 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(9)·1870-8; Tugwell P, et al. Clinical improvement as reflected in measures of function and health-related quality of life following treatment with leflunomide compared with methotrexate in patients with rheumatoid arthritis. *Arthritis Rheum* 2000,43(3):506-514

85 Strand V, 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(9)·1870-8. The SF-36 has proved to be valid and reliable in a large number of diseases in addition to RA (e.g., cardiovascular disease, low back pain, Type II diabetes, and osteoarthritis) Ware JE, et al. The MOS 36-item short form health survey (SF-36). *Medical Care* 1992;30(6)473-483; Ware JE, et al. SF-36® Health Survey. Manual and Interpretation Guide. Lincoln, RI: Quality Metric Incorporated, 1993, 2000; Ware JE and Kosinski M SF-36 Physical & Mental Health Summary Scales: A Manual for Users of Version 1, 2nd ed Lincoln, RI: Quality Metric, 2001.

86 Strand V, 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(9).1870-8; Tugwell P, et al. Clinical improvement as reflected in measures of function and health-related quality of life following treatment with leflunomide compared with methotrexate in patients with rheumatoid arthritis *Arthritis Rheum* 2000,43(3):506-514.

physical impairment and reductions in health related quality of life in RA patients. *See supra,* fns. 80 and 82.

3. Slowing Of Radiographic Progression

HRG also fails to note that Arava® significantly retarded or inhibited progression of RA as shown by radiographic evidence in both of the placebo-controlled trials.[87] This occurred at 12 months in US301 and at 6 months in MN301. In both trials, Arava® reduced the progression of structural joint damage by more than 75% compared to placebo.[88] In US301 and MN302, the slowing of progression was comparable for Arava® and methotrexate, with no consistent difference across the two studies.[89] These data are comparable to those reported for the other recently approved DMARDs. [90]

4. The Benefits Of Arava® Were Maintained In A Second Year Of Treatment

Double-blind treatment was continued to 24 months in the US301 trial and the extensions of the MN301 and MN302 trials. These 2-year data were published and available to HRG -- but ignored.[91] These data confirmed that clinical efficacy in Arava®-treated patients was sustained over 2 years of treatment. The benefits achieved during the first year of Arava® treatment -- reduction in signs and symptoms, improvements in physical function, and the slowing or inhibiting radiographic progression -- were maintained in patients continuing a second year of treatment.[92]

---

87 As measured by total Sharp scores, which sum (add) erosions and joint space narrowing.

88 Of the placebo patients in US301, approximately 60% had received active treatment in an alternative therapy phase for a mean of 7 to 8 months after withdrawing from placebo treatment.

89 In the two trials comparing Arava® and methotrexate, the slowing of radiographic progression was statistically significant in favor of Arava® in US301 (p=0.0499), and the two drugs were not statistically different in MN302, demonstrating overall similar effect. Of interest, US301 was also the first placebo-controlled trial to demonstrate the efficacy of methotrexate in slowing radiographic progression. Likewise, MN301was the second placebo-controlled trial to demonstrate efficacy of sulfasalazine in slowing radiographic progression, and the slowing of progression with sulfasalazine was statistically equivalent to Arava® (p=0 3394). Sharp JT, et al. Treatment with leflunomide slows radiographic progression of rheumatoid arthritis *Arthritis Rheum* 2000;43(3).495-505

90 Strand V, Sharp JT: Review· Radiographic Data from Recent randomized controlled trials in RA: What have we learned? *Arth Rheum* 2002; 46: (accepted for publication).

91 Cohen S, et al Two-year, blinded, randomized, controlled trial of treatment of active rheumatoid arthritis with leflunomide compared with methotrexate. *Arthritis Rheum* 2001,44(9)·1984-1992; Scott DL, et al. Treatment of active rheumatoid arthritis with leflunomide· two year follow up a double blind, placebo controlled trial versus sulfasalazine. *Ann Rheum Dis.* 2001,60:913-923, Kalden JR, et al. Improved functional ability in pateints with rheumatoid arthritis—long-term treatment with leflunomide versus sulfasalazine. *J Rheum* 2001;28(9): 1983-91, Emery P, et al A comparison of the efficacy and safety of leflunomide and methotrexate for the treatment of rheumatoid arthritis. *Rheumatology* 2000;39·655-665.

92 *Id*

III. THE PETITION MISREPRESENTS THE SAFETY OF ARAVA®

In addition to mischaracterizing the efficacy data, HRG posits a selective and misleading review of the clinical and post-marketing safety surveillance data. In fact, the clinical studies confirmed that Arava® is safe and effective when used according to the prescribing information, and nothing in the post-marketing experience contradicts that conclusion.

A. CLINICAL TRIALS ESTABLISHED THE SAFETY OF ARAVA®

The FDA's determination of Arava®'s safety was based upon an integrated clinical trial database containing safety data from over 2400 patients in phase II and III studies, including over 1300 rheumatoid arthritis patients receiving Arava ®. This database also represents the largest blinded, controlled exposure for methotrexate therapy in RA.

EXPOSURE in Phase II and III Clinical Trials:

Treatment Group	Total Exposed	≥ 6 Months	≥12 Months	I year data Patient Years	2 year data Patient Years
LEF	1,339	1,011	838	2077	2467
MTX	680	549	497	936	1558
SSZ	133	76	23	258	244
PL	310	90	38	226	256

More than 800 Arava® patients were in the phase III studies alone. At the time the Arava® NDA was filed with the FDA, it was the largest database ever submitted for approval of a DMARD in RA. The 12-month primary safety analysis of the three phase III clinical trials was provided in detail in the Arava® NDA, and 2 year integrated safety data were thereafter provided to the FDA.[93]

Notwithstanding these substantial safety data, HRG refers to only limited results that appear to skew the safety analysis. For example, HRG suggests that "[in] assessing hepatotoxicity, the most weight . . . should be given to US301," in which folate (which reduces not only side effects such as liver enzyme elevations, but may also reduce efficacy of

---

93 The phase III clinical trials, including the 2 year data from these trials, has been published Strand V, et al. Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate. *Archives Int Med* 1999, 159·2542-2550; Smolen JS, et al Efficacy and safety of leflunomide compared with placebo and sulphasalazine in active rheumatoid arthritis· a double blind, randomized, multicentre trial. *Lancet* 1999;353.259-66, Cohen S, et al. Two-year, blinded, randomized, controlled trial of treatment of active rheumatoid arthritis with leflunomide compared with methotrexate. *Arthritis Rheum* 2001,44(9)·1984-1992; Scott DL, et al. Treatment of active rheumatoid arthritis with leflunomide: two year follow up a double blind, placebo controlled trial versus sulfasalazine. *Ann Rheum Dis.* 2001;60.913-923; Kalden JR, et al. Improved functional ability in pateints with rheumatoid arthritis—long-term treatment with leflunomide versus sulfasalazine. *J Rheum* 2001,28(9) 1983-91; Emery P, et al. A comparison of the efficacy and safety of leflunomide and methotrexate for the treatment of rheumatoid arthritis. *Rheumatology* 2000;39.655-665.

methotrexate) was required. *See* Petition at 3. HRG does not mention the results of MN302 (in which folate was not required and was taken by only 10 percent of methotrexate patients), where the incidence of adverse liver events with methotrexate was significantly higher than with Arava®, yet the incidence in Arava® treated patients in MN302 was comparable to US301. Later in the Petition, however, HRG reverses its position on the importance of US301 and disregards it, suggesting instead that MN302 establishes "superior" efficacy. *See* Petition at 16. An assessment of drug safety should not be based on selective and contradictory use of the same data.

As discussed below, not only was HRG selective in its use of data, but even the information cited does not support its position. The clinical trial safety data -- the very basis for the FDA's approval of the Arava® NDA -- have not changed since they were submitted to the FDA. The data supported the FDA's conclusion that Arava® could be safely used when it was first approved in 1998, and it still supports that conclusion.

1. The Frequency And Severity Of Adverse Events Involving Arava® Were Similar To Those With Methotrexate And Sulfasalazine

HRG selectively relies on data from one trial (US301) to suggest that patients treated with Arava® experienced adverse events of greater frequency and severity than those associated with the active comparator drugs. For example, HRG claims that more Arava® patients withdrew due to adverse events compared to methotrexate. In fact, the rate of withdrawal from US301 for serious adverse events was the same for Arava® and methotrexate, and the total number of treatment-related serious adverse events (as judged by the investigators, not the sponsor) was less with Arava®. The FDA mandated withdrawals for asymptomatic elevated LFTs. *See* Appendix B, Table 3. A clearer understanding of safety emerges from a review of the integrated adverse event data from all phase III clinical trials that were provided to FDA, as well as data from individual trials that were published but ignored by HRG:

- Serious adverse events[94] occurred in similar numbers of Arava® and methotrexate patients (and slightly less with sulfasalazine). Fewer Arava® patients had serious

94 The term "serious adverse events" is defined by the Code of Federal Regulations to include any adverse drug experience occurring at any dose that results in any of the following outcomes: Death, a life-threatening adverse drug experience, inpatient hospitalization or prolongation of existing hospitalization, a persistent or significant disability/ incapacity, or a congenital anomaly/birth defect. Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered a serious adverse drug experience when, based upon appropriate medical judgment, they may jeopardize the patient or subject and may require medical or surgical intervention to prevent one of the outcomes listed in this definition Examples of such medical events include allergic bronchospasm requiring intensive treatment in an

adverse events assessed by the investigator as treatment-related as compared to both methotrexate and sulfasalazine across all studies.[95] *See* Appendix B, Table 2.

- Serious adverse events considered by the investigator to be treatment-related and withdrawals due to treatment-related serious adverse events were less frequent with Arava® than with methotrexate plus folate in US301.[96]

- The treatment-related serious adverse events in the Arava® and placebo groups in US301 consisted of 1 patient in each group with asymptomatic LFT elevations not requiring hospitalization and 1 patient in each group with non-fatal sepsis. In contrast, the treatment-related serious adverse events in the methotrexate group consisted of 2 patients with asymptomatic LFT elevations not requiring hospitalization, 1 patient with pneumonia, 1 patient with interstitial pneumonitis, and 1 patient with fatal sepsis.[97]

- The year-2 incidence of serious adverse events for the year-2 cohort was similar across treatment groups (leflunomide = 25.3%; sulfasalazine = 26.7%; methotrexate= 20.8 in US301and 27.2% in MN304).[98]

- Serious adverse events in year 2 assessed by the investigator as possibly treatment-related were similar among the Arava® and both methotrexate groups, and fewer than the sulfasalazine group.[99]

- In year-2, there were fewer withdrawals for all adverse events, including fewer withdrawals for serious adverse events and treatment-related serious adverse events in the Arava®-treated patients than in either of the methotrexate groups and fewer than in the sulfasalazine group.[100]

- Deaths occurred at a similar rate among the treatment groups in year 1 and year 2 of the phase III controlled trials. In the first year of the three phase III studies, death occurred

---

emergency room or at home, blood dyscrasias or convulsions that do not result in inpatient hospitalization, or the development of drug dependency or drug abuse. *21 CFR 314.80.*

95 *See* Appendix B, Table 2, which provides an overview of the Adverse Events (AEs) reported in the phase II (leflunomide patients only) and phase III clinical trials in the 1 year database of NDA 20-905.

96 Strand V, et al. Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate. *Archives Int Med* 1999,159:2542-2550. The authors state that serious adverse events assessed as treatment-related by the investigator were reported for 2 patients receiving Arava® (1 1%), 2 patients receiving placebo (1.7%), and 5 patients receiving methotrexate (2.7%). *See also* Appendix B, Table 3.

97 *Id.*

98 *See* Appendix B, Table 4. Of the patients in those studies who completed 12 months of treatment, 450 Arava®-treated patients entered a second year of double-blind treatment in US301, extension MN305, or extension MN304. In US301, 101 patients treated with methotrexate with folate continued for a second year. From MN302, 320 patients treated with methotrexate without folate entered the MN304 second year extension. From MN301/303, 60 sulfasalazine entered the MN 305 second-year extension. The patients who entered a second year of double-blind treatment were designated the "year-2 cohort" and were evaluated in a supplemental integrated analysis of safety in the second year of treatment The 2- year safety analysis compared safety in the year-2 cohort second year of treatment to the first year of treatment in the same patients. In turn, year-1 of the year-2 cohort was compared to year 1 of the intent-to-treat (ITT) population (i.e., all patients randomized to receive at least one dose of study drug in phase III trials and extension). The leflunomide treatment groups from the three phase III trials were pooled for the supplemental 2-year safety analysis Methotrexate treatment groups were not pooled because folate was required in US301 whereas only 10 percent of methotrexate patients in MN 302/304 received folate. The supplemental 2-year safety data analysis also included an additional 8 leflunomide patients and 8 methotrexate patients from 5 Canadian sites that were not a part of the primary 1 year data analysis because Canada joined the US301 study a year after the other sites.

99 *Id.*

100 *Id*

in 0.7% of Arava®-treated patients which was similar to the rate in the sulfasalazine (0.8%), methotrexate without folate (1.2%), and methotrexate with folate (0.5%)groups.

- In year-2 of treatment, death occurred in 0.7% of Arava®-treated patients, which was less than in both of the methotrexate treatment groups (1.0% for methotrexate with folate in US301 and 2.2% for methotrexate without folate in MN302). No deaths occurred in the 60 sulfasalazine patients in year-2.

- Similar proportions of Arava® and methotrexate patients withdrew due to adverse events, and more withdrew on sulfasalazine, in the phase II and III NDA studies. *See* Appendix B, Table 2.

- Fewer adverse events assessed by the investigator as treatment-related, and less dose reduction due to adverse events, occurred in Arava®-treated patients than in the methotrexate or sulfasalazine patients in the phase II and III NDA studies . *Id.*

2. HRG Mischaracterized The Adverse Event Profile Of Arava® For Several Disease Endpoints

HRG focuses on certain adverse events (while selectively ignoring others) that occurred during the clinical trials. For example, HRG mentions vasculitis and suggests that Aventis failed to report two clinical trial deaths associated with vasculitis.[101] This accusation is false and misleading. First, the eventual deaths of these two patients occurred long after they withdrew from the clinical trial, as is clear from the publication on which HRG relies. Second, both cases were reported to the FDA during the trial at the time the vasculitis was diagnosed, and both were detailed in the NDA submission. Third, both deaths were, in fact, reported by Aventis to the FDA after the trials were concluded.[102] Moreover, vasculitis is listed in the **Adverse Reactions: Cardiovascular** section of the prescribing information,[103] based on occurrence in the phase II and III clinical trials at a rate of 0.6% with Arava®, which was similar to methotrexate (0.6%) and to sulfasalazine (0.8%).[104]

---

101 HRG refers to a letter to the editor in 1999 describing two patients who withdrew from MN302 due to vasculitis and who subsequently died 10 months and more than 2 years later, respectively *See* Petition, at p.13, citing Bruyn GAW, et al. Leflunomide for active rheumatoid arthritis. *Lancet* 1999; 353:1883.

102 *Id* Although neither of the eventual deaths occurred during the trial or the post-trial observation period, follow-up information was available at the time of the NDA submission (regarding the male patient who died 10 months after withdrawal from the trial) and was forwarded to the FDA in addition to being included in the NDA The follow-up information regarding the other patient (a female who died more than 2 years after withdrawal from the trial) was reported to the FDA when the information became available through the publication to which HRG refers

103 Arava® prescribing information.

104 Smolen JS, et al. Reply to Bruyn GAW et al. Leflunomide for active rheumatoid arthritis. *Lancet* 1999, 353: 1883-1884.

a. Hypertension

HRG is correct that hypertension was reported more often in Arava® patients than in the control groups, but HRG only tells half the story. Of the Arava® patients with hypertension, a significant proportion (ranging in the phase III clinical trials from 75% to 100%) had evidence of pre-existing hypertension, either from a diagnosis of hypertension at study entry or hypertensive blood pressure readings at baseline. The incidence of new-onset hypertension was low, and there was no significant difference among treatment groups.[105] Moreover, the potential causal impact of concomitant NSAID and steroid use could not be excluded, as all subjects with new onset hypertension were receiving one or both of those drugs.

b. Hepatic Events

Detailed analyses of liver enzyme elevations in the phase III studies of Arava® were provided in the NDA submission, including incidence and degree of elevation of both hepatic aminotransferases -- alanine aminotransferase (ALT) and aspartate aminotransferase (AST).[106]

HRG is correct that, in US301, mild elevations occurred more often in patients treated with Arava® (17.6%) than in patients treated with methotrexate with folate (11.0%).[107] However, HRG disregards the fact that the incidence of clinically significant elevations (>2xULN)[108] and the subset of marked elevations (>3xULN) in Arava®-treated patients was similar to the methotrexate with folate group in US301and much less than the methotrexate without folate group in MN302.[109] These clinically significant elevations – both moderate (>2 to ≤ 3xULN) and marked (> 3xULN) -- in Arava® patients were generally reversible while continuing treatment or with dose reduction or discontinuation.[110]

HRG also focuses on two patients in the phase III clinical trials who had ALT elevations of 39xULN and 80xULN respectively, but fails to note that both cases were detailed in the NDA submission and the etiologies for both were confounded by other factors, as assessed by the

---

105 These ranged from 0 to 2.2% in the Arava® groups, 0 to 1 1% in the placebo groups, 0.4 to 1.6% in the methotrexate groups, and 0 8% in the sulfasalazine group.

106 Because ALT is more sensitive to elevation than AST, and because patients in the studies with AST elevation also had ALT elevation (generally to a higher level), ALT elevations are shown in the appendix. ALT elevations are categorized based on the highest elevation for an individual patient. *See* Appendix B, Table 5

107 Mild ALT elevations (>1.2 to ≤2x ULN) occurred in 14 4% to 17 6% of Arava®-treated patients across the phase III trials, *see id.*, Table 5, and 98 percent of these normalized to < 1.2xULN generally while continuing treatment.

108 ULN = Upper Limits of Normal.

109 *See* Appendix B, Table 6; Arava® prescribing information

110 *See* Appendix B, Table 6. For all clinically significant elevations (>2x ULN), there was no difference in normalization rate with Arava (49/59, 83%) and methotrexate (148/174, 85%). Additionally, when all ALT elevations (> 1.2x ULN} are considered, the normalization

FDA.[111] It is important to note that in the NDA, 14 other cases of severe (>8xULN) ALT elevations did not involve Arava®: 1 in a placebo-treated patient, 2 in sulfasalazine-treated patients, and 11 in methotrexate-treated patients.

In short, the phase III clinical trials showed that the incidence of clinically significant liver enzyme elevations in Arava®-treated patients was similar to the incidence in patients on methotrexate with folate and lower than in patients on methotrexate without folate. Most ALT elevations were mild, and elevations were generally asymptomatic and reversible. Furthermore, the incidence of these events during the second year of Arava® treatment was not higher than during the first year of treatment, indicating that incidence does not increase with extended duration of treatment.[112] Accordingly, there is no basis for concluding that the clinical trials evidence any greater risk of hepatotoxity, as defined by elevated LFTs, in Arava® patients compared to methotrexate patients.

c. Lymphoma

HRG suggests without basis that Arava® is associated with an increased risk of lymphoma. However, in the clinical trial data in the Arava® NDA submission, the overall incidence of malignancies did not substantially differ between treatment groups, including placebo. Various malignancies were reported in all groups, but frequencies were low and there was no clustering of findings in particular organs. Furthermore, the 2-year data for the active treatment groups did not demonstrate a higher incidence of malignancy for Arava®.

Rheumatoid arthritis is believed to be associated with an increased risk of lymphoproliferative disorders. In the absence of any clinical trial evidence of increased incidence of malignancy in Arava® patients, but based on the known increased risk of lymphoproliferative disorders associated with the use of some immunosuppressive medications, the **Warnings** section in the prescribing information clearly states:

> Malignancy
> The risk of malignancy, particularly lymphoproliferative disorders, is increased with the use of some immunosuppression medications. There is a potential for

---

rate was higher in Arava-treated patients than in methotrexate-treated patients 173/186 (93%) for Arava compared to 243/278 (87%) for methotrexate.

111 Both patients were taking other drugs associated with hepatic events (one was taking sustained release niacin and lovastatin and the other was taking diclofenac with pre-existing Hepatitis C infection and had recently tapered her own prednisone dose, without knowledge of her treating physician. Both patients discontinued leflunomide treatment, with cholestyramine washout; and liver enzyme elevations resolved once the other drugs associated with potential hepatic toxicity were discontinued.

112 In year-2, both ALT elevations and abnormal LFTs reported as adverse events occurred with lower frequency compared to year-1 as shown in Appendix B, Table 7. Two patients had ALT elevations >3x ULN that had not reversed to <2x ULN at the end of the study, but they subsequently reversed on follow-up.

immunosuppression with ARAVA. No apparent increase in the incidence of malignancies and lymphoproliferative disorders was reported in the clinical trials of ARAVA, but larger and longer-term studies would be needed to determine whether there is an increased risk of malignancy or lymphoproliferative disorders with ARAVA.

d. Other Serious Adverse Events Of Interest With DMARD Therapies

The controlled phase III trials provided no evidence that other adverse events of a serious nature that are considered related to DMARD therapies occurred more frequently with Arava® than with methotrexate or sulfasalazine treatment. There were no cases of interstitial pneumonitis, renal failure, or agranulocytosis in the Arava®-treated patients (representing 1333 patient years of exposure over 2 years of treatment in the phase III studies and 2467 patient years in the combined phase II and III clinical trials over two years), although these events were seen in the methotrexate and sulfasalazine control groups in the same phase III clinical trials over the same time period with less drug exposure (i.e., fewer patients exposed and fewer patient years of exposure).[113]

3. The Year-2 Clinical Trial Safety Data Are Consistent With the Year-1 Data

The adverse event profile of Arava®during the second year of treatment was similar to that during the first year of treatment, with no new types of adverse events emerging. The incidence of liver enzyme elevations decreased in the second year of treatment. Long-term information on the safety of therapy over 2 years supports its continued tolerability without emergence of new patterns of adverse events, either serious or non-serious, and with a diminished overall adverse event rate in a second year of treatment.

Based on the clinical data, there is no basis for concluding that methotrexate or sulfasalazine are "safer" than Arava®. To the contrary, analysis of the safety data from the controlled phase III studies shows that the overall percentage of patients with treatment-related serious adverse events and withdrawals due to serious adverse events (treatment related or not) was generally similar with Arava®, methotrexate, and sulfasalazine administration.

---

113 For example, there were two cases of agranulocytosis in the sulfasalazine patient group. Sulfasalazine also had the highest incidence of lymphoproliferative disorders. In the methotrexate-treated patients, there was renal failure, as well as four cases of interstitial pneumonitis (one of which was fatal) and one case of interstitial fibrosis. The rate of nonfatal sepsis was higher in the methotrexate groups than in the Arava® group. No Arava® patients developed pancytopenia, whereas pancytopenia in a methotrexate patient led to fatal pneumonia. The incidence of vasculitis was similar among the treatment groups in the clinical trials, and all were less than 1 per 100 patient years, rheumatoid vasculitis is a known extra-articular manifestation of RA.

B. HRG MISCHARACTERIZES THE POST-MARKETING ADVERSE EVENT PROFILE OF ARAVA®

As with all other treatments for RA -- and most other prescription drug products -- adverse events have been reported in association with the post-marketing use of Arava®. Rather than reviewing those reports objectively in the context of the disease state and itsassociated morbidities, background incidence of certain events, polypharmacy (multiple medications) and the presence or lack of confounding factors, HRG offers yet another selective and inaccurate interpretation of the data.[114] As discussed below, an objective review of the post-marketing data confirms that there is no factual basis to conclude that the risk profile for Arava® is less favorable than that of other available DMARD therapies. This Response will address the various categories of adverse events mentioned by HRG in the Petition.

1. Limitations Of Post-Marketing Data

In evaluating post-marketing data, it is important to understand the limitations of "spontaneous" reports and the purpose of reviewing such data. Spontaneously reported post-marketing information is evaluated with regard to potential new adverse health consequences and/or an increased incidence or severity of known risks.[115] The number of cases reported may vary considerably depending on the treatment; comparisons with other agents or estimated background rates of events in a given disease are difficult. However, the likelihood of under-reporting is lower with a newer drug such as Arava® than with other established, widely used therapies, such as methotrexate. Under-reporting is more likely with an older drug, such as methotrexate (used for 25 years and formally approved for RA in 1986).[116] Other factors that may affect the reporting of adverse events include: novelty of the event; severity of the event; perceived relationship to drug administration; adverse effects reported with similar drugs; physician awareness; previous reports of an adverse reaction (either in clinical trials or post-marketing surveillance data); and media interest.[117]

---

114 The search that HRG conducted using the FDA's AERs database was not exhaustive and did not capture all events reported for either Arava® or methotrexate. For most disease endpoints identified in the Petition, more adverse events were reported for methotrexate than for Arava®

115 PhRMA/FDA/AASLD Drug-Induced Hepatotoxicity White Paper – Postmarketing Considerations, November 2000 (the "White Paper"), p 3

116 Tsong Y, Comparing reporting rates of adverse events between drugs with adjustment for year of marketing and secular trends in total reporting, *J of Biopharm Stat*, 1995; 5(1). 95-114

117 White Paper, p.3.

Another recognized limitation of spontaneous reporting is that many events have no more than a temporal association with the use of a drug, and differential reporting can make the benefit-risk profile of two drugs appear very different when, in fact, they are not.[118] As noted in a recent PhRMA/FDA/AASLD White Paper, " [w]ith the exception of some drug-specific diseases or symptoms . . . , the risk in unexposed patients (background risk) is never zero, so that reports of a drug association may be incorrect, and instead reflecting only background occurrence of the event."[119] As stated in the FDA's MedWatch form, anecdotal case reports do not establish causation – this is particularly the case in a disease with well recognized co-morbidities

The following discussion addresses HRG's distorted review of the post-marketing data.

2. Post-Marketing Reports Of Hepatic Events.

Analysis of spontaneously reported hepatic events requires objective consideration of several factors, none of which appear to have been addressed by HRG. First, concomitant use of Arava® with other treatments for RA, including DMARDs, in addition to other confounding factors, make determination of a causal relationship between Arava® and any given event uncertain. For example, methotrexate, sulfasalazine, gold, azathioprine and cyclosporine have all been associated with hepatic events. Second, because RA is a systemic disease that can affect many extra-articular organs, underlying disease activity must also be considered as a potential causal factor, in addition to frequent co-morbid conditions such as cardiovascular disease. Third, hepatic events have been reported with other drugs, including both prescription and non-prescription drugs used in the treatment of RA. When complete information is lacking, as is often the case with post-marketing surveillance data, it is difficult to determine whether any or all of these potential contributing factors may be responsible for the adverse events reported following Arava® use.

Aventis has applied standardized case definitions and criteria for assessing causation with respect to all serious and non-serious reports of hepatic events from post-marketing clinical trials and post marketing surveillance.

---

118 *Id* (Emphasis added).
119 *Id.*

a. **September 1998 to September 2001**

During the 3-year period from September 1998 to September 2001, Aventis received 126 reports of adverse hepatic events (including serious and non-serious) that were classified as possibly associated with Arava® use, utilizing criteria for definition, classification, and analytical methods described by an international panel of experts for drug-induced hepatotoxicity.[120] The majority of these cases were classified as hepatocellular, with some cholestatic or mixed pattern events.

Of all hepatic adverse event reports, 23 were associated with a fatal outcome where *any* hepatic event was reported; a fatal hepatic event was specifically reported in 11 of the 23 cases. In the remaining 12 of these 23 cases, liver abnormalities were only one of several events in patients with multiple morbidities, and were not reported to be the cause of the fatal outcome.

In order to better understand these fatal events, Aventis consulted an outside expert, Professor Dominique Larrey, of the Hepatology and Transplant Unit, School of Medicine, Montpellier, France, to review the 23 cases in detail. Dr. Larrey concluded that none of the cases exhibited a definite causal relationship to Arava®administration; and that Arava® possibly could have had a contributory role in six of the reported cases due to the temporal relationship between Arava® use and the event. He considered the data to be consistent with a rare potential for hepatotoxicity, based primarily on the increases in ALT and the number of hepatic events reported.[121]

It is generally recognized that accurate incidence rates for adverse events cannot be established from spontaneous post-marketing surveillance data due to the absence of a certain and defined denominator (the total number of patients who were prescribed the treatment and complied with the prescription), as well as the variable degree of reporting adverse events, influenced to some degree by the perceived or documented safety profile of the agent at the time of its approval; specific adverse event labeling, and the well recognized degree of under-reporting inherent in a spontaneous reporting system. Furthermore, the nature of a voluntary reporting system often results in collection of incomplete information, and subsequent follow-up reports may be confused and counted as new events. Nonetheless, reporting rates may be roughly

---

120 Benichou C, Danan G Causality assessment of Adverse Reactions to Drugs-I. A novel method based on the conclusions of international consensus meetings: application to drug-induced liver injuries. *Journal Clinical Epidemiology* 1993;46:1323-1330.

121 The 1/21/02 Expert Report of Professor Dominique Larrey has previously been provided to the FDA.

estimated using as a denominator the number of patient years of exposure calculated from product sales information.

Based on the available data, an estimated overall reporting rate for fatal hepatic events (11 cases) is 5.7/100,000 patient years. As noted, after application of internationally recognized case definitions and causality criteria, as well as analysis by an external expert hepatologist, a possible association was assessed in 6 of the 11 fatal hepatic cases.[122] The estimated reporting rate for these 6 possible fatal hepatic reactions is 2.25/100,000 patient years. As a point of reference, the occurrence rate of fatal hepatic events in the general population has been estimated by EMEA to be 11.7/100,000 patient years.[123]

b. September 2001 to March 2002

During the 6 month period from September 2001 through March 2002, Aventis received 24 reports of adverse hepatic events that were classified as possibly associated with Arava® administration, using the definitions, classifications, and analytical methods identified above. Distribution according to the type of liver injury reflects the same profile as in the previous three-year period, with a predominant hepatocellular pattern. In addition to these 24 cases, there were three cases where a hepatic event (liver failure) was reported as the fatal event. Of these three cases, however, none was assessed as possibly related to Arava® therapy: in one case, autopsy revealed hepatitis B infection; the second case was confounded by multiple concomitant medications; and the third case lacked any clinical information for assessment.[124]

Since first marketed, the prescribing information for Arava® has contained information about potential hepatotoxicity in the **Warnings** section, including monitoring recommendations. The rare serious hepatic events observed in the post-marketing period do not alter the positive benefit-risk profile of Arava®.[125]

3. Post-Marketing Reports Of Lymphoma

Aventis has received 13 spontaneous reports of lymphoma from 1998 to March 2002. In 5 of the 13 cases, the reporting physician assessed the event as unrelated to Arava® therapy. In

---

122 *Id.*

123 EMEA Benfit-Risk Assessment for Arava® (available to FDA upon request). Exposure data represents the three years from September 1998 through September 2001. If all 23 cases were considered to be causally related, the occurrence rate would be 11.9/100,000 patient years.

124 These cases have not yet been reviewed by Dr Larrey

125 As previously noted, Aventis is working with the FDA to update the prescribing information to include additional data regarding the rare serious post-marketing hepatic events reported in in association with Arava®.

those cases where sufficient information was available (11 out of 13 cases), excluding one case of interrupted therapy, symptoms that led to the diagnosis of lymphoma occurred between 2 and 6 months after the first Arava® dose. Occurrence of malignancy after such short-term exposure is considered an unlikely case for drug-induced pathology.[126] Most patients had received concomitant or previous long-standing treatment with other DMARDs, including methotrexate; persistent active RA and prolonged use of immunomodulatory treatments, such as DMARDs, are associated with a greater risk for lymphoma in patients with RA.[127] In addition, methotrexate has been associated with lymphomas that occurred during treatment and regressed upon discontinuation of this therapy.[128] The post-marketing reports in patients taking Arava® have not demonstrated such a pattern.

In the general population in 1997, the age-adjusted incidence rate of lymphoma was 15.8 per 100,000. The 1993-1997 age-adjusted incidence rate was 16.0 per 100,000. An increased incidence of lymphoma and/or lymphoproliferative disorders is believed to be associated with the underlying inflammatory RA disease process.[129]

Assuming as a worst case analysis -- that all 13 reported cases were causally associated with Arava® -- the observed reporting rate in Arava®-treated patients would be approximately 4.9 cases of lymphoma per 100,000 patient years -- lower than the estimated incidence rate of lymphoma in the general population. Post-marketing case reports of lymphoma therefore do not suggest evidence of a new safety signal;[130] as previously stated, the prescribing information includes a warning regarding this potential risk. *See supra,* Section III.A.2.c.

4. Post-Marketing Reports Of Hematologic Events.

It is difficult to interpret many reports of hematologic events because of: (i) hematologic abnormalities associated with RA; (ii) use of other medicines associated with hematologic adverse events; and (iii) pre-existing conditions in some patients. For example, methotrexate and sulfasalazine are associated with severe and sometimes fatal hematologic events. Nevertheless, on February 23, 2000, the **Warnings, Precautions**, and **Adverse Reactions** sections of the

126 Cancer Principles and Practice of Oncology, *6th ed, 2001.*

127 *Id.*; ACR Hotline: FDA Advisory Committee reviews safety of TNF inhibitors, ACR 9/24/01. *See also* Appendix C, List of 70 additional lymphoma/RA references.

128 Genovese M Musculoskeletal Syndromes in Malignancy In: Kelley's Textbook of Rheumatology, 6th edition, Ruddy S et al, eds., WB Saunders Co, Phila 2001; Weinblatt M. Methotrexate. In: Kelley's Textbook of Rheumatology, 6th edition, Ruddy S et al, eds., WB Saunders Co, Phila 2001.

129 Ries CAF, Eisner MP, Kosary CL,et al. SEER Cancer Statistics Review, 1993-1997, National Cancer Institute. NIH Pb. No. 00-2789 Bethesda, MD 2000

Arava® prescribing information were amended (following FDA approval) to inform physicians that there had been rare spontaneous reports of pancytopenia in patients receiving Arava®. The prescribing information also included the following statement:

> In most cases, patients received concomitant treatment with methotrexate or other immunosuppressive agents, or they had recently discontinued these therapies; in some cases, patients had a prior history of a significant hematologic abnormality. If ARAVA is used in such patients, it should be administered with caution and with frequent clinical and hematologic monitoring.

Aventis communicated these labeling changes to health care providers in a Dear Doctor letter dated March 21, 2000, which indicated the need to monitor for hematologic effects when used in combination with other hematotoxic DMARDs, which also require hematologic monitoring.

In the **Warnings** section of the prescribing information, hematologic monitoring is recommended for patients at increased risk of hematologic toxicity. In the **Adverse Reactions** section of the approved prescribing information, thrombocytopenia, leukopenia, and anemia are listed. On February 23, 2000, this section was amended (following FDA approval) to include post-marketing events of pancytopenia. Moreover, the **Warnings** section was also amended at that time to state that Arava® is not recommended in patients with severe immunodeficiency, bone marrow dysplasia or severe uncontrolled infections.[131]

Accordingly, the post-marketing data do not provide evidence of a greater risk of hematologic events than what is already referenced in the prescribing information.

5. Post-Marketing Reports Of Dermatologic Events

On February 23, 2000, following receipt by Aventis of reports of Stevens-Johnson syndrome ("SJS") and toxic epidermal necrolysis ("TEN"), the **Warnings** and **Adverse Reactions** sections of the prescribing information were amended (following FDA approval) to include SJS and TEN as well as erythema multiforme to inform physicians of the occurrence of these rare events and to provide recommendations for the drug elimination procedure. Aventis communicated these changes in labeling to health care providers in a Dear Doctor letter dated March 21, 2000.

---

130 Use of the term "signal" does not mean that a finding of causation between the drug and the event(s) has been established; rather, the term refers to surveillance information that suggests a need to conduct additional evaluation and/or analysis.

131 Aventis has also been working with the FDA to update the hematologic monitoring recommendations contained in the prescribing information.

The number of reports for these events has remained stable since launch, despite increased exposure to Arava®. In many of these reported cases, confounding factors, including concomitant medications such as antibiotics and NSAIDs, which are also associated with these severe skin reactions, were present.

Neither the Petition nor the post-marketing surveillance data provide evidence of significantly greater risk of severe dermatologic events with Arava® than with other DMARDs. The reports remain rare and are adequately described in the prescribing information.

6. Post-Marketing Reports Of Hypertension

HRG claims that physicians are uninformed about the risk of hypertension because the prescribing information does not mention hypertension as a post-marketing adverse event. This argument is specious. Table 5 in the prescribing information identifies adverse events occurring in 3 percent or greater of clinical trial patients; hypertension is specifically mentioned under the heading "**Cardiovascular.**"

7. Post-Marketing Reports Of Pregnancy

HRG does not claim that post-marketing data require withdrawal of Arava®. Instead, HRG briefly discusses the pre-clinical toxicology data, but makes no specific recommendation. These data, as well as half-life of the active metabolite and elimination process to remove any effect of active drug were extensively discussed at the FDA Arthritis Advisory Committee hearing. The resulting recommendations are reflected in the label and include a washout procedure using 8 g of cholestyramine 3 timesper day for 11 days (representing conservative estimates regarding blood levels and half life of the active metabolite), as well as two blood level determinations indicating no active drug (or metabolite) prior to pregnancy (offered by the sponsor upon request without cost to the patient).[132]

HRG also notes reports of post-marketing experience of maternal exposure to Arava®, concluding that "safe" levels of maternal exposure are unknown. This topic was covered in great detail in the FDA Arthritis Advisory Committee hearing, and, further, the boxed warning at the

---

132 The effectiveness of the cholestyramine washout procedure was tested in phase 1 trials, in addition to the one study referenced by HRG HRG also distorts the safety profile of Arava® in its discussion of half-life and elimination. First, the Petition mistakenly suggests that a long half-life makes Arava® inherently unsafe. This is simply wrong. There are drugs with very short half-lives that can be unsafe and drugs with long half-lives that are safe. Second, the Petition suggests that since Arava has a long half-life, it may be stored somewhere in the body and have negative effects a long time after discontinuation. This is again false. There is no pharmacokinetic evidence for storage or compartmentalization of Arava or its metabolites anywhere in the body. Instead, Arava's long half-life is due to interohepatic recycling in the liver, which sends the active metabolite from the liver to the bile and from the bile to the GI tract, where it is reabsorbed into the body. In turn,

beginning of the Arava® prescribing information expressly states that "Pregnancy must be excluded before the start of treatment . . . Arava® is contraindicated in pregnant women, or women of childbearing potential who are not using reliable contraception. . . . Pregnancy must be avoided during Arava® treatment . . . ."[133]

Aventis nevertheless continues to evaluate the clinical impact of exposure during pregnancy and is sponsoring a multi-center cohort study established by the Organization of Teratology Information Services (OTIS). The program provides counseling as well as post-marketing surveillance relative to the potential teratogenicity of Arava®. The study will document pregnancy outcome with respect to the presence or absence of a pattern of malformation in liveborn infants in women with first-trimester prenatal exposure to Arava®. Secondary endpoints to be evaluated include the rate of spontaneous abortions or stillbirth, pre- or post-natal growth deficiency, and premature delivery.[134]

8. Post-Marketing Reports Of Gastrointestinal Events

HRG does not suggest that Arava® should be withdrawn due to post-marketing reports of severe diarrhea. Instead, HRG notes that more reports were identified for Arava® than for methotrexate. Based on data from the controlled clinical trials, it is not surprising that more post-marketing reports of GI events were received with respect to Arava® treatment, because the incidence in these clinical studies was higher in patients receiving Arava® compared with those receiving metrotrexate.

Neither the character nor frequency of post-marketing surveillance adverse events indicate a greater risk of gastrointestinal events with Arava treatment than was observed in the clinical trials, and described in detail in the product label.

---

cholestyramine enhances elimination of Arava® by interrupting (and preventing) the interohepatic recycling process in the GI tract, where the active metabolite binds with the cholestyramine, thus preventing reabsorbtion into the body

133 HRG's allegation that evidence of minimal maternal exposure proves that the warning is ineffective is equally specious. Under this theory, no product with any warnings of serious adverse events before or during pregnancy should be marketed. Moreover, HRG offers no evidence suggesting that the physicians in the reported cases were unaware of the warning.

134 HRG questions the effectiveness of the wash-out procedure to remove or reduce plasma levels of the active metabolite. The drug elimination procedure in the prescribing information is designed to achieve nondetectable plasma levels of <0.2 mg/L (0 02 μg/ml). This level is more than 136 and 123 times lower, respectively, than Cmax levels in rats and rabbits, which did not cause embryotoxicity or teratogenicity Dr. Robert Brent, a leading teratologist and FDA consultant, notes in a recent article (that HRG cites but ignores on this point, *see* Petition at 14) that a 100-fold reduction represents a conservative approach. Brent, RL Teratogen Update. Reproductive Risks of Leflunomide (Arava); A pyrimidine Synthesis Inhibitor· Counseling Women Taking Leflunomide Before or During Pregnancy and Men Taking Leflunomide Who are Contemplating Fathering a Child, *Teratology* 2001, 63. 106-112 Finally, it should be noted that post-washout M1 (active metabolite) plasma levels were not detected in 97 percent of the post-marketing reports of washout. This is convincing evidence of the effectiveness of the cholestyramine washout procedure.

9. Post-Marketing Reports Of Weight Loss

HRG refers to weight loss as an adverse event reported more frequently in Arava® than methotrexate treated patients. Not only is it difficult to relate weight loss in individual patients with administration of Arava® (or other DMARDs), but a unified mechanism to explain these observations is lacking. Based on limited observations in the phase II trials with Arava® treatment, the phase III randomized controlled trials specifically included physical and laboratory evaluations when clinically significant weight loss was observed. Across all phase III trials, few reports of treatment-associated weight loss required these pre-specified, additional analyses. Mean changes in weight, lipid profiles and other parameters, including serum total protein and albumin levels, in the leflunomide groups compared with placebo or active comparators failed to identify treatment associated changes.

Although it is difficult to evaluate post marketing surveillance reports of treatment-associated weight loss, it is likely that multiple etiologies explain these observations. Although patients with poorly controlled, active RA frequently complain of fatigue and malaise associated with elevated IL-6 levels, increased production of pro-inflammatory cytokines including TNFa and IL-1 in active rheumatoid arthritis result in profound systemic manifestations of malaise and fatigue, characterized as an anorectic/catabolic state. Weight loss may therefore result from organic (gastrointestinal disorders, connective tissue disease, endocrine, infection, malignancy, pulmonary, and neurologic), psychological and/or idiopathic etiologies. To determine whether reported weight loss is due to the underlying inflammation of rheumatoid arthritis or its treatment may not only be difficult but, in fact, impossible.[135] Anecdotal reports of weight loss as well as weight gain, and positive as well as negative changes in lipid profiles have occurred with other recently approved biologic and synthetic DMARDs.[136] To date, it has not been possible to ascertain whether these changes are treatment related or clinically meaningful.

When other gastrointestinal symptoms, including anorexia, nausea, vomiting, and diarrhea are reported, weight loss may reflect treatment associated adverse events. With Arava treatment, reports of weight loss have not included either baseline bodyweights or the period of time when weight loss was observed/reported. Nor were relevant clinical data provided, making it virtually impossible to establish a treatment associated causal relationship.

---

135 Cope AP. Regulation of autoimmunity by proinflammatory cytokines *Curr Opin Immunol* 1998; 10:669-76.

Of the five case reports cited by HRG,[137] one patient discontinued Arava® and initiated etanercept therapy, and reportedly weight remained stable after discontinuation of Arava®. The four other patients continued on Arava® treatment due to good clinical responses, and their weight stabilized after the initial self-limited reports of weight loss. Based on the above analysis, the post-marketing reports do not reflect any new signal or an increased frequency or severity of weight loss. Accordingly, the information in the prescribing information adequately warns of a potential for weight loss in association with the use of Arava®.

* * * * * *

For the reasons stated above, post-marketing surveillance data do not represent a reliable comparison between a recently approved treatment such as Arava® and the standard of care, methotrexate, used for the past 25-30 years and specifically approved for the treatment of RA in 1986. It is important to remember that concern regarding LFT elevations with methotrexate therapy remain; specific guidelines for monitoring treatment have facilitated broad utilization in RA without requiring liver biopsies prior to treatment initiation and at intervals thereafter.[138] Familiarity with methotrexate therapy without requiring pre- and interim-treatment liver biopsies, has evolved over 16-25 years of clinical use, indicating that rheumatologists will carefully monitor DMARD therapies for active RA, recognizing they offer significant clinical benefits, but are nonetheless associated with significant potential risks. These treatments require detailed knowledge of the underlying autoimmune disease and careful monitoring of its therapy.

As a conservative estimate, RA patients have at least 30-40 years of active disease, and will need more treatments than are currently available to remain physically active and able to engage in work and social activities they deem important. Arava, as well as other recently approved DMARDs, represents a significant addition to the therapeutic armamentarium. However, even if a patient had the best and most prolonged clinical response to each of these therapies (as predicted by the clinical trials), used in a conservative, sequential fashion, they will not be sufficient in treating this lifelong debilitating disease with its associated co-morbidities.

---

136 Vis M, Nurmohamed MT, Wolbink G et al. Short term effects of infliximab on lipid profiles in patients with RA Ann Rheum Ds 2002, 61.S75.

137 Coblyn JS, et al. Leflunomide - Associated weight loss in rheumatoid arthritis. *Arthritis Rheum* 2001; 44(5):1048-1051.

138 Kremer JM, et al. Methotrexate for rheumatoid arthritis. suggested guidelines for monitoring liver toxicity *Arthritis Rheum* 1994;37(3):316-328; ACR Ad Hoc Committee on clinical guidelines Guidelines for monitoring drug therapy in Rheumatoid Arthritis. *Arthritis Rheum* 1996,39(5)·723-731.

Despite the approval of 7 new treatments for RA (4 new DMARDs and 3 COX-2 inhibitors) in the last 4 years, this disease still represents a significant unmet clinical need.

IV. THE BENEFITS OF ARAVA® OUTWEIGH ASSOCIATED RISKS

A substantive benefit-risk analysis for a RA treatment must be based on a clear understanding of the underlying disease and available therapies, as well as a thorough evaluation of safety and efficacy data. Rather than offering a reasoned, scientific evaluation, HRG cites clinical and post-marketing data without providing appropriate context. This unbalanced and selective approach does a disservice to the many thousands of patients who benefit from Arava® therapy.

Rheumatoid arthritis is a serious, crippling disease, with a high personal and socio-economic cost. There is no known cure. The risks and benefits of Arava® must be evaluated in the context of the manifestations and severity of the disease, and the strengths and limitations of other available therapies. All DMARDs have efficacy in the treatment of RA -- and all are associated with serious adverse events and require careful clinical and laboratory monitoring. A wide choice of DMARDs is needed in clinical practice to address issues of tolerability and decreased efficacy over time, especially in a disease that may last for 20 to 30 years, or more. Arava® has a unique mechanism of action that prevents the production of T-cells through the inhibition of pyrimidine synthesis -- targeting the disease process of RA. As demonstrated herein, a comprehensive analysis of the data compels the conclusion that the benefits of Arava® therapy outweigh known risks.

These conclusions are reinforced by two recent studies. One, a placebo-controlled study, confirms significant efficacy of Arava® when used in combination with methotrexate. The second, a 40,000 patient retrospective cohort study, shows that Arava®-treated patients generally had fewer adverse events overall than patients taking methotrexate or other DMARDs.

A. THE PLACEBO-CONTROLLED STUDY OF COMBINATION ARAVA® AND METHOTREXATE SUPPORTS THE POSITIVE BENEFIT-RISK PROFILE FOR ARAVA®

Arava® and methotrexate have different mechanisms of action -- inhibition of pyrimidine synthesis (Arava®) versus inhibition of intracellular purine pathways of metabolism resulting in modulation of cytokine and adenosine levels (methotrexate) -- which suggests a potential for

benefit in combination through complementary actions, especially in patients with inadequate response to monotherapy with either drug.

US4001 was a phase IIIb (post-marketing) study of combination Arava® and methotrexate.[139] The study evaluated the efficacy and safety of adding Arava® in RA patients inadequately responding (with active disease) to methotrexate, as compared to adding placebo to methotrexate.[140] It was a 6-month, multi-center trial involving 263 patients that was placebo-controlled, randomized, and double-blind. At the end of the 6-month study, patients were allowed to enter an open label extension phase for an additional 6 months. Patients on placebo were switched to Arava® at that time without using a loading dose. During the open-label phase, patients remained blinded to their original randomized treatment arm.[141]

1. Efficacy Results

US4001 demonstrated the efficacy of adding Arava® in RA patients who had active disease while on methotrexate alone.[142] The ACR20 Responder-at-Endpoint rate after adding Arava® (46%) was more than twice that after adding placebo (20%). When Arava® was added to ongoing methotrexate therapy, more than half of these patients (52%) were ACR 20 Responders at their last study visit compared to 23% receiving placebo. One-half of the Arava®-treated patients who were ACR 20 responders were also ACR 50 responders (at least 50% improvement). The ACR 50 and ACR 70 (at least 50% and 70% improvement, respectively) responder rates for Arava® were statistically significantly higher than placebo rates. The substantial benefit was also observed with regard to physical function. HAQ Disability Index improved significantly with the addition of Arava® compared to the addition of placebo. US4001 has provided additional support for the efficacy of Arava® compared to

---

139 Kremer JM, et al. Concomitant leflunomide therapy in patients with active rheumatoid arthritis despite stable doses of methotrexate: A randomized comparison of efficacy, safety, and tolerability compared to methotrexate alone. *Annals Int Med* 2002 (accepted for publication); Kremer JM, et al. The combination of leflunomide and methotrexate in patients with active rheumatoid arthritis who are failing on MTX treatment alone: a double-blind placebo controlled study *Arthritis Rheum* 2000: 43(9):S224; Furst DE, et al. Adding leflunomide to patients with active rheumatoid arthritis while receiving methotrexate improves physical function and health-related quality of life *Arthritis Rheum* 2000; 43(9). S224.

140 It should be noted that study 4001 was not a comparison of Arava plus methotrexate combination therapy versus methotrexate monotherapy; rather, it was a comparison of Arava versus placebo when added to background MTX in patients with persistent active disease while on methotrexate alone. These were patients who were selected for tolerating MTX monotherapy without LFT elevation Therefore, the patients randomized to adding placebo would be expected to have a low incidence of LFT elevations, which was in fact the case.

141 Kremer JM, et al. Concomitant leflunomide therapy in patients with active rheumatoid arthritis despite stable doses of methotrexate: A randomized comparison of efficacy, safety, and tolerability compared to methotrexate alone. *Annals Int Med* 2002 (accepted for publication); Kremer JM, et al. The combination of leflunomide and methotrexate in patients with active rheumatoid arthritis who are failing on MTX treatment alone. a double-blind placebo controlled study. *Arthritis Rheum* 2000. 43(9):S224, Furst DE, et al. Adding leflunomide to patients with active rheumatoid arthritis while receiving methotrexate improves physical function and health-related quality of life. *Arthritis Rheum* 2000; 43(9): S224.

142 *Id.*

placebo, in this case when added to background methotrexate treatment, demonstrating that patients who are inadequately responding to methotrexate can achieve clinically and statistically meaningful improvement by adding Arava®.

2. Safety Results

The safety findings of this phase IIIb combination therapy study -- the type and frequency of adverse events -- were consistent with those reported in clinical trials evaluating Arava® monotherapy. No clinical hepatic adverse events (i.e., a clinical diagnosis, as opposed to laboratory abnormalities alone) were reported during the 6-month, placebo-controlled study or the subsequent 6-month, open-label extension. Analysis of laboratory values showed that most of the ALT and AST elevations were mild (≤ 2 x ULN), as they were in the phase III monotherapy studies. The incidence of clinically significant (>2 x ULN) ALT elevation and the subset of marked ALT elevations (>3 x ULN) after adding Arava® to ongoing methotrexate was within the range observed with Arava® monotherapy in the phase III trials. The highest ALT elevation was 4.8xULN.[143]

HRG assumes, based solely on a study report of one patient with liver cirrhosis confounded by many years of methotrexate treatment (which is associated with cirrhosis), that "the temptation to combine leflunomide and methotrexate holds many dangers." *See* Petition at 6. US4001 demonstrated that adding a lower initial dose of Arava® than is recommended for monotherapy,[144] with subsequent increase or decrease as appropriate for the individual, allowed the combination to be used effectively with a safety profile consistent with that seen in the phase III monotherapy studies of Arava®. Aventis currently is in discussion with the FDA regarding the addition of information relating to this study to the prescribing information.[145]

---

143 *See* Appendix B, Table 8.

144 A lower dose than recommended for Arava monotherapy was used. The loading dose 100 mg daily for 2 days, rather than 3 days, and the initial maintenance dose was 10 mg daily, rather than 20 mg, which was adjusted upward or downward as necessary.

145 Additionally, Aventis has recently completed Study HWA/486/4002. This multinational study was designed to evaluate whether the combination of leflunomide and sulfasalazine was superior to sulfasalazine alone, for the treatment of active RA, in patients who were non-responders after 24 weeks of leflunomide. Dosing levels contained in the U.S. prescribing information for monotherapy were used for a 6 month open label period, at the conclusion of which non-responders to leflunomide monotherapy were randomized to either sulfasalazine or placebo. Of the 968 patients initially treated with leflunomide, only 106 patients were non-responders who advanced to the second phase of the trial, which was such a small sample size that no meaningful comparisons could be drawn. In short, there was a substantially higher than expected response to leflunomide monotherapy (672 patients).

B. THE COHORT STUDY

The relative safety of Arava® is further supported by the results of a retrospective cohort study of more than 40,000 RA patients.[146] Aventis used the claims database of a large managed care organization and compared the rates of liver, blood, skin, hypertension, and other adverse events among users of Arava®, alone and in combination with other DMARDs, to rates among users of methotrexate and other DMARDs, alone and in combination. The cohort of patients mirrored the larger RA population within the United States in terms of age, sex, and drug treatment. It is the largest cohort study of DMARD therapies in RA patients involving head-to-head comparisons of DMARDs.[147]

The results show that Arava® monotherapy is associated with fewer adverse events overall (12.20 AEs per 100 patient years) than other DMARDs, including methotrexate (18.85). Arava® monotherapy is also associated with a statistically significantly lower incidence of hypertension and respiratory events than other DMARD monotherapies, including methotrexate. The incidence of adverse events for other outcomes (hepatic, hematologic events, skin disorders, and pancreatitis) were not statistically different (though the rates were lower) than the other DMARDs.[148] Moreover, the combination of Arava® and methotrexate had significantly lower overall incidence of adverse events than the two comparator combinations (leflunomide plus other DMARDs and methotrexate plus other DMARDs). Finally, the mortality rate among Arava® users was lower than the comparison groups (there was one death in the Arava® group, 9 in the methotrexate group, and 82 in the DMARD group); these rates, however, were not statistically different. These results are shown in the following table, and the rates shown are reported per 100 patient years (except for mortality, where the rates are per 100,000 persons). This table also captures the total patient years for each DMARD or DMARD combination.

---

146 Cannon GW, Holden WL, Hochberg M, Juhaeri J, Dai W, Scarazzini L, Stang P. Adverse Events with Disease Modifying Antirheumatic Drugs. a Cohort Study With Comparison of Leflunomide with other DMARDs (submitted for publication) In addition to the Cohort study, Aventis performed five additional epidemiologic analyses of the available data, which were provided to the EMEA/CPMP and to the FDA, all of which confirm the positive benefit-risk profile for Arava®

147 The cohort study design allowed for the determination of person-time exposure of individuals, i.e., the time (in years) that a person is at risk for the development of a particular adverse event (the denominator), and whether that person actually had the event (if yes, the numerator). The resulting incidence rate -- the numerator divided by the denominator -- can be further adjusted for the potential confounding effects of age, sex, and other medical conditions, which may distort the true association between drug use and adverse event

148 Arava® monotherapy had a higher, though not significantly different, incidence of hematologic events than methotrexate (0 14 per 100 PY vs. 0 08 per 100 PY).

	LEF (2166 PY)	MTX (4808 PY)	Other DMARD (15717 PY)	NSAID (7028 PY)	COX-2 (3894 PY)	LEF + MTX 1024 PY)	LEF + other DMARD (2719 PY)	Other DMARD + MTX (8621 PY)
Any AE	12.20	18.85	18.89	40.37	33.78	5.31	7.40	8.55
Hepatic	0.45	0.70	0.58	1.35	1.07	0.53	0.24	0.34
Hematologic*	0.14	0.08	0.24	0.20	0.13	n/c	0.04	0.10
Skin*	n/c	0.12	0.10	0.02	n/c	n/c	0.04	0.03
Hypertension	3.98	6.65	6.10	16.77	14.18	1.68	2.47	2.75
Pancreatitis*	0.24	0.25	0.33	0.53	0.10	0.14	0.18	0.16
Respiratory	2.40	5.26	4.84	9.21	7.69	1.71	1.62	2.31
Mortality**	121.9	279.6	469.5	92.0	n/a	145.2	201.5	156.2

*For hemotologic, skin, and pancreatitis, there were too few events or too little person-time for the mathematical model to adjust for age, sex, and comorbidities in all exposure groups

**Rates per 100,000 persons

As noted above, the data for the Cohort Study came from a managed care organization claims database. Limitations of such a database include lack of indicators of disease severity, limited clinical detail, little or no data on compliance and use of over-the-counter drugs, as well as patient history. Nevertheless, the data are consistent with the conclusion that Arava® has a safety profile similar to the other DMARDs, including methotrexate. To be sure, HRG has offered no valid analysis to the contrary.

V. THE STANDARD FOR WITHDRAWAL CANNOT BE MET

Arava® (leflunomide) Tablets is a "new drug" as defined under section 201(p) of the Federal Food, Drug & Cosmetic Act ("FFDCA"), 21 USC 321, and it is the subject of an approved NDA, 21 USC 355. Following approval of a NDA, the Secretary is authorized to withdraw approval of a new drug only under limited circumstances (pursuant to the section 505(e) of the FFDCA) and only after giving due notice and an opportunity for hearing to the applicant. In order to withdraw an application, the Secretary must determine that at least one of the following facts is present:

1. clinical or other experience, tests, or other scientific data show that a drug is unsafe for use under the conditions of use that formed the basis for approval of the application;

2. new evidence of clinical experience evaluated together with the evidence available when the application was approved, shows that the drug is not shown to be safe for use under the conditions of use that formed the basis for approval of the application; or

3. new information evaluated together with the evidence available when the drug was approved, shows that there is a lack of substantial evidence that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling.[149]

Moreover, HRG has requested the Secretary to "immediately remove" Arava® from the market. The only authority to do so is "if the Secretary finds that there is an imminent hazard to the public health, he may suspend the approval" of a NDA immediately.[150] This extraordinary action may be undertaken "only in the exceptional case of an emergency, which does not permit the Secretary to correct it by other means."[151]

As demonstrated above, HRG has failed to prove any of the bases for withdrawal:

1. Arava® is not "unsafe" and has a safety profile comparable to other available DMARDs;

2. There is no new evidence of clinical experience warranting withdrawal; and

3. There is no new information that suggests that Arava® does not have the effect it purports to have.

Indeed, the overwhelming weight of the data -- including the most recent clinical and other information -- provides further evidence of the positive benefit-risk profile of Arava®. The Petition, therefore, is unsupportable and must be denied.

VI. CONCLUSION

The benefit-risk profile of Arava® remains positive, and nothing HRG has submitted demonstrates otherwise. RA is a severe, chronic and disabling disease with no known cure. The arsenal of therapies available to treat RA is limited, and all of them have certain drawbacks. Unfortunately, there is no panacea for treating RA, and no single DMARD is effective for all

149 21 USC 355(e)
150 *Id* This authority cannot be delegated
151 Sen. Rep. No. 1744 at 7, 87th Cong., 2d Sess. (1962).

patients throughout the course of their disease. Arava® is an important option available to physicians who treat patients with RA, as reflected in the unsolicited letter submitted by Dr. Gary S. Firestein, M.D., Chair of the Arthritis Advisory Committee, in opposition to the Petition. *See* Appendix A.

The randomized, controlled, phase III clinical trials demonstrate that Arava® is both safe and effective when used in accordance with the FDA approved prescribing information, and nothing in the post-marketing experience suggests otherwise. Thus, the legal standard applicable to the withdrawal of an NDA has not been met by HRG, and the Petition should be denied.

SUBMITTED BY:

AVENTIS PHARMACEUTICALS INC.
August 8, 2002

UNIVERSITY OF CALIFORNIA, SAN DIEGO UCSD

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

DIVISION OF RHEUMATOLOGY, ALLERGY AND IMMUNOLOGY
SCHOOL OF MEDICINE
9500 Gilman Drive
La Jolla, Ca 92093-0656

Tel (858) 534-2359
Fax: (858) 534-2606

GARY S FIRESTEIN, M D
Professor of Medicine and Chief

June 10, 2002

Food and Drug Administration
Washington, D.C.

To whom it may concern,

A recent Citizen's Petition was submitted to the Department of Health and Human Services regarding the safety of leflunomide. The authors requested that this drug be withdrawn from the market due to its toxicity. In light of the importance of these issues and the need place the petition's comments into perspective, I would like to offer my unsolicited opinion on the matter. As the chairman of the FDA Arthritis Advisory Committee, a practicing physician/rheumatologist for over 20 years, a translational researcher on the pathogenesis of rheumatoid arthritis (RA), and the executive director of a clinical trial center (cit.ucsd.edu), I believe that I can provide some insights that will be useful to the FDA. I should note that the specific details of individual patient histories are not available to me, and that my conclusions are based on the information provided in the petition and my own familiarity with the field.

The first issue that needs to be considered when evaluating the safety of any treatment for RA is that toxicity must be compared with the morbidity and mortality associated with active inflammatory synovitis. RA is not a benign condition, and many studies have demonstrated significantly higher mortality compared with controls (reviewed in Br J Rheumatol 1993;32 Suppl 1:28-37). This is especially true for patients with significant limitations on their activities of daily living, evidence of active inflammatory disease (e.g., high CRP), or involvement of many joints. While the impact of treatment on mortality is not fully understood, recent information suggests that effective treatment can prolong life (Lancet 2002; 359:1173-7). The mechanism of improved survival is not established, but is probably directly related to suppression of synovial and systemic inflammation. The impact of active RA on quality of life also needs to be considered when evaluating the risk/benefit ratio of a therapeutic agent. In other words, merely describing the potential toxicity of an agent in a vacuum is not only insufficient but can be misleading.

Because of the serious long-term consequences of active RA, rheumatologists have become increasing aggressive in its management. Immunosuppressive agents, cytokine antagonists, anti-metabolites, and combination therapy have become mainstays. Instead of relying on the now outdated "pyramid" approach, treatment is initiated early and is accelerated rapidly in order to suppress inflammation (Am J Med. 2001;111:498-500). Clinical trials using aggressive management, such as the COBRA trial and many others, have demonstrated improved outcomes compared with conservative approaches. In this context, the conservative and risk-averse recommendations of the Citizen's Petition clearly fail to take into account two key elements of modern management: 1) poorly controlled RA is a dangerous and morbid condition; and 2) aggressive treatment can alter the natural history of the disease.

With regard to some of the specific toxicity issues raised in the document, one can stipulate that leflunomide can be hepatotoxic. However, the information provided in the petition does not accurately

address either the risk/benefit ratio or how the drug fits into the constellation of agents available for use in RA. For instance, there are a variety of assertions regarding the relative safety of methotrexate compared with leflunomide. Perhaps most important is the putatively lower rate of hepatotoxicity of the former. The comparative data are not derived from controlled databases, but from voluntary physician reporting. There is a well-described bias introduced when comparing toxicity of established agents to new agents that is clearly evident in this analysis. There is also little information on the use of concomitant drugs or the assiduousness of monitoring that could have prevented serious adverse events. Therefore, it is impossible to draw a conclusion regarding the relative rates of serious adverse events based on this information. The comments related to the long half-life of leflunomide raise reasonable concerns; however, clinical practice has supported the adequacy of cholestyramine in many cases where toxicity has been observed. Based on the data provided by the petition, it would be appropriate to recommend a study of the relative toxicities of methotrexate and leflunomide in a more controlled setting. However, withdrawing an effective agent like leflunomide based on this limited information is both unjustified and counterproductive.

Perhaps the most important consideration in this discussion is how leflunomide should be used compared with other anti-rheumatic agents. Even if one assumes that methotrexate is a safer agent, current clinical practice guidelines indicate that leflunomide should be primarily administered to patients that have an inadequate response to methotrexate or have other contraindications. This makes comparisons of the relative toxicities moot, since patients that receive leflunomide would, by definition, have active disease and already received a putatively safer agent. Since we already know that active RA is an unacceptable alternative, then we are obliged to advance therapy using agents that are either less effective, more toxic, or have other undesirable attributes (e.g., expense or requirement for parenteral administration).

The alternatives to leflunomide suggested in the petition under these circumstances do not accurately represent state-of-the-art clinical practice. For instance, the use of "Rest and nutrition" as recommended by the Merk Manual is part of the outdated pyramid approach that does not recognize the long-term consequences of active RA. Of the "slow acting" agents recommended, two (gold and penicillamine) have not been used by most rheumatologist for over a decade due lack of efficacy and toxicity that far exceeds leflunomide. Hydroxychloroquine and especially sulfasalazine are stated to be equivalent to methotrexate and leflunomide. Sulfasalazine has been used extensively to treat patients with RA, especially in Europe. However, clinical experience in the United States does not support the assertion that it is as effective as methotrexate or leflunomide. The reported equivalence with sulfasalazine is likely due to inadequate dosing of comparators or type II errors due to underpowered studies. Immunosuppressive agents, including cyclosporine and azathioprine, have considerable toxicity and limited efficacy. Reliance on a tertiary source like the Cochrane Library or the Merck Manual as in the petition to determine the relative efficacy does not necessarily provide the most up to date or useful information.

Overall, patients that have an inadequate response to methotrexate are typically treated with a TNF inhibitor, leflunomide, or sulfasalazine (either alone or, more commonly, in combination). The selection of a particular agent depends on the patient's particular circumstances. Moreover, the percentage that respond to each of these drugs is limited, which means that several might be tried to determine the optimum combination. For instance, only 15% of patients failing methotrexate that receive the TNF inhibitors have an ACR70 response and only about 30% achieve an ACR50 response. The response rates for sulfasalazine are likely lower. Therefore, most patients will require considerable experimentation to find the best combination of drugs. Removing one of these key agents from our armamentarium would be a major setback to their management and is unjustified.

The final comments in the petition relate to the ineffectiveness of changing labels or educating physicians. On the contrary, the dissemination of information through the physician and patient

community is now rapid and has high penetration. For instance, new guidelines to assess patients receiving TNF inhibitors for prior tuberculosis exposure had a major impact on clinician practice. The rapidity of processing new information is especially true for RA because new anti-rheumatic drugs are mainly prescribed by subspecialists. The notion that rheumatologists do not modify their practice after appropriate education is simply untrue and is likely based on outdated information. The influence of patient advocacy also should not be underestimated. In my own clinical practice, the majority of patients receiving leflunomide specifically asked about the safety issue.

In conclusion, vigilance in post-marketing safety is a major concern and one must be ready to act if appropriate signals are observed. In the case of leflunomide, one must be cognizant of the risks of uncontrolled RA, the relative lack of efficacy for the alternatives to methotrexate, and the contribution of inadequate monitoring or inappropriate combination therapy to severe reactions. Leflunomide is an effective agent in RA that decreases inflammation, improves quality of life, and slows the progression of disease. The information provided by the petition does raise questions that should be addressed with appropriate studies, and the concomitant use with methotrexate should be carefully addressed. However, withdrawing the agent is simply not justified with the current information and would lead to increased morbidity (and possibly mortality) in RA patients that do not respond to methotrexate.

Sincerely,

Gary S. Firestein, M.D.
Professor of Medicine
UCSD School of Medicine

Chairman
FDA Arthritis Advisory Committee

B

APPENDIX B
CLINICAL TRIAL EFFICACY AND SAFETY TABLES

Efficacy results from the Phase III clinical trials and the US 4001 study of combination Arava plus methotrexate are provided in Table 1 showing ACR response rates, HAQ Disability Index which measures physical function, and total Sharp scores which measure x-ray progression.

Table 1. Efficacy Results in Clinical Trials (Intent-to-treat Analysis)[1]

Study# Design	Pts at BL	Treatment Group (n)	ACR≥20% Responder-at-Endpoint[3] (% of pts)	ACR Responder[2] rates (LOCF)			HAQ Disability Index[4]		Sharp score (xray)[5]	
				ACR≥20% (% of pts)	ACR≥50% (% of pts)	ACR≥70% (% of pts)	Mean BL	Mean change	Mean BL	Mean change
US 301 12 mo PC, R, DB	482	LEF (182)	41[a]	52[a]	34[af]	20[ae]	1.30	-0.45[ae]	23.11	0.53[ad]
		PLA (118)	19	26	8	4	1.31	0.03	25.37	2.16
		MTX (180)	35[a]	46[a]	23[a]	9	1.30	-0.26[d]	22.76	0.89[c]
MN 301[6] 6 mo PC, R, DB	358	LEF (130)	49[b]	55[a]	33[b]	10[c]	1.65	-0.56[ag]	46.26	1.23[a]
		PLA (92)	29	29	14	2	1.59	-0.08	46.18	5.88
		SSZ (132)	45[c]	57[a]	30[b]	8	1.50	-0.37[d]	41.86	2.32[d]
MN 302 12 mo R, DB	999	LEF (501)	43	51	31	10	1.50	-0.44	24.94	2.48
		MTX (498)	57[e]	65[e]	44[e]	16[e]	1.52	-0.54[g]	24.60	1.62
US 4001 6 mo PC, R, DB	263	Ongoing MTX +LEF (130)	46[a]	52[a]	26[a]	10[c]	1.6	-0.42[a]	n.d.	n.d.
		+PLA (133)	20	23	6	2	1.5	-0.09	n.d.	n.d.

R = randomized; PC = placebo controlled; BL=baseline; LOCF=last observation carried forward; n.d.=not done; LEF=leflunomide; MTX= methotrexate; PLA=placebo; SSZ=sulfasalazine

[1] Intent-to-treat (ITT) population defined as all patients randomized who received at least one dose of study drug with at least 1 study evaluation. ITT subjects who did not have an evaluation after randomization (leflunomide 3, methotrexate 2, sulfasalizine 1) were not in the efficacy analysis but were in the safety analysis.

[2] An ACR 20 Responder is defined by the ACR as a patient who had 20% or greater improvement in 5 of 7 core set measures of disease activity [Felson A&R 1995]. An ACR 20 Responder may also fulfill criteria for higher thresholds of response; an ACR 50 or ACR 70 Responder is defined in an analogous manner to the ACR 20 Responder, but with improvements of at least 50% or 70%, respectively.

[3] An "ACR 20 Responder-at-Endpoint" is a patient who completed the study and was an ACR 20 Responder at the completion of the study. (Any patient discontinuing early was counted as a nonresponder.)

[4] HAQ=Health Assessment Questionnaire Disability Index (Score 0=Best, 3=Worst). A decrease in score indicates improvement.

[5] Retardation of structural damage compared to control was assessed using the Sharp Score [Sharp, JT. Scoring Radiographic Abnormalities in Rheumatoid Arthritis, *Radiologic Clinics of North America*, 1996; vol. 34, pp. 233-241], a composite score of erosions and joint space narrowing in hands/wrists and forefeet.

[6] In the publication [Smolen et al Lancet 1999], ACR20 Responder-at-Endpoint rates are given as LEF 48% and SSZ 44%, and ACR20 Responder rate for SSZ is given as 56%.

Arava or MTX or SSZ vs. placebo: [a]$p \leq 0.001$; [b]$p \leq 0.01$; [c]$p \leq 0.02$; [d]$p \leq 0.05$
Arava vs. MTX or SSZ: [e]$p \leq 0.01$; [f]$p \leq 0.02$; [g]$p \leq 0.05$.

[Strand V, et al. Treatment of Active Rheumatoid Arthritis with leflunomide compared with placebo and methotrexate. *Arch Intern Med* 1999;159:2542-2550; Smolen JS, et al. Efficacy and safety of leflunomide compared with placebo and sulphasalazine in active rheumatoid arthritis: a double blind, randomized, multicentre trial. *Lancet* 1999;353:259-66; Kalden JR, et al. Improved functional Ability in Pateints with Rheumatoid Arthritis—longterm treatment with leflunomide versus sulfasalazine. J Rheum 2001;28(9): 1983-91; Emery P, et al. A comparison of the efficacy and safety of leflunomide and methotrexate for the treatment of rheumatoid arthritis. *Rheumatology* 2000;39:655-665; Sharp JT, et al. Treatment with leflunomide slows radiographic progression of rheumatoid arthritis. Arthritis Rheum 2000;43(3):495-505; Kremer JM, et al. *Annals Int Med* 2002 (in press); Kremer JM, et al. The combination of leflunomide and methotrexate in patients with active rheumatoid arthritis who are failing on MTX treatment alone: a double-blind placebo controlled study. *Arthritis Rheum* 2000: 43(9):S224; Furst DE, et al. Adding leflunomide to patients with active rheumatoid arthritis while receiving methotrexzte improves physical function and health-related quality of life. *Arthritis Rheum* 2000; 43(9): S224; Arava (leflunomide) Prescribing InformationTable 2 provides an overview of the Adverse Events (AEs) reported in the Phase II and III clinical trials in the 1 year database of NDA 20-905. The Arava (LEF) group includes all rheumatoid arthritis patients in the Phase II and III trials. Placebo (PLA), methotrexate (MTX), and sulfasalazine (SSZ) groups are those of the Phase III controlled trials.

Table 2. Overall Summary of Adverse Events: Phase II/III NDA trials				
	LEF (n=1339)	**PLA n=210)**	**MTX (n=680)**	**SSZ (n=133)**
	%	%	%	%
Subjects w/ 1 or more AE	83.4	82.9	92.9	91.0
Subjects w/ 1 or more drug related AE	59.8	51.4	69.7	73.7
Subjects reducing dose due to AE	4.0	0	14.1	6.8
Subjects discontinuing due to AE	15.5	7.1	13.4	22.6
Subjects w/ 1 or more SAE	22.0	10.5	21.9	16.5
Subjects w/ 1 or more drug-Related SAE	4.9	3.3	6.3	8.3

AE= adverse event; SAE= serious adverse event

In Table 3, the Phase III adverse events leading to withdrawal and serious adverse events are summarized by study and by treatment groups within each study.

Table 3. Phase III Clinical Trials: Adverse Event Withdrawals and Serious Adverse Events (% of patients)

	301US (12 months)			301MN (6 months)			302MN (12 months)	
	LEF (182)	PL (118)	MTX (182)	LEF (133)	PL (92)	SSZ (133)	LEF (501)	MTX (498)
All AE withdrawals	22.0	8.5	10.4	14.3	6.5	18.8	18.8	14.9
Due to LFTs	7.1	1.7	4.4*	1.5	1.1	1.5	1.6	3.2
SAEs	14.8	10.2	8.2	17.3	13.0	13.5	31.1	26.9
Related	1.1	1.7	2.7	5.3	5.4	6.8	7.2	7.6
Withdrawals	3.3	1.7	3.3	5.3	3.3	3.0	9.0	5.6
Related	1.1	0.8	2.2	3.8	2.2	3.0	3.8	3.0
LFTs	0.5	0.8	1.1	0	0	0.8	0.2	0.6

LFT=liver function test

* In the methotrexate group of the US 301 study, there were a total of 8 patients (4.4%) who withdrew due to LFT adverse events as in the study report and summary tables in the NDA Briefing Document section 6.5.3.2 and in the published manuscript [Strand V et al. *Arch Int Med* 1999; 159:2542-2550]. These include 2 patients who withdrew due to an adverse event reported as SGPT (ALT) increased and/or SGOT (AST) increased in addition to the 6 patients (3.3%) cited by HRG who withdrew due to an adverse event reported as LFT abnormal.

Table 4 summarizes the year-2 incidences of adverse event withdrawals and serious adverse events for the year-2 cohort treatment groups of the Phase III studies.

Table 4. Adverse events leading to withdrawal and serious adverse events with onset in year-2
Phase III studies: year-2 cohorts

	% of patients			
	LEF (N=450)	**SSZ (N=60)**	**MTX US301 (with folate) (N=101)**	**MTX 304 (without folate) (N=320)**
All AEs leading to withdrawal	4.0	13.3	7.9	4.4
SAEs	25.3	26.7	20.8	27.2
Related	3.1	8.3	2.0	1.6
Withdrawal	0.9	5.0	6.9	1.6
Related	0.4	1.7	2.0	0.6

Detailed analyses of liver enzyme elevations in the three Phase III studies of Arava monotherapy were provided in the NDA submission, including incidence and degree of elevation of both of the hepatic transaminases, alanine aminotransferase (ALT) and aspartate aminotransferase (AST). ALT is more sensitive to elevation than AST with more frequent and higher elevations, and patients in the studies with AST elevation also had ALT elevation, generally to a higher level. For that reason, ALT elevations are shown in the following tables.

Table 5 –Percent of patients with ALT elevations in Phase III monotherapy trials of leflunomide: Categorized by highest elevation

	LEF			MTX		PLA		SSZ
	US301 %	MN301/3† %	MN302 %	US301 %	MN302 %	US301 %	MN301 %	MN301/3† %
ALT >1.2 to ≤2xULN	17.6	17.3	14.4	11.0	16.9	6.8	10.9	10.5
ALT >2.0 to ≤3.0xULN	6.6	1.5	4.4	6.6	14.9	0	0	5.3
ALT >3.0xULN	4.4	1.5	2.6	2.7	16.7	2.5	1.1	1.5
Total ALT >1.2xULN	28.6	18.8	21.4	20.3	48.4	9.3	12.0	14.3
Total ALT > 2.0xULN	11.0	2.3	7.0	9.3	31.5	2.5	1.1	6.0

ULN = Upper limit of normal range.

† Includes MN303 extending the data to 12 months for the active treatment arms.

Liver enzyme elevations were generally reversible while continuing treatment or with dose reduction or discontinuation. Reversibility of clinically significant (>2xULN) ALT elevations is shown in Table 8. The table provides the number that reversed to <2x ULN and also the number that normalized to <1.2x ULN. It also states whether the normalization occurred after drug discontinuation for any reason, after dose reduction, or after no change in dose.

Table 6. Reversibility ALT elevations: Phase III trials

ALT (SGPT)	US301[1] (12 mos)			MN301/303[2] (12 mos)			MN302[3] (12 mos)	
	LEF	PLA	MTX	LEF	PLA	SSZ	LEF	MTX
>3-fold ULN n(%)	8 (4.4)	3 (2.5)	5 (2.7)	2 (1.5)	1 (1.1)	2 (1.5)	13(2.6)	83 (16.7)
Reversed to ≤2-xULN	8	3	5	2	1	2	12‡	82
Normalized to ≤ 1.2xULN	7	3	5	2	1	1	9	73
after discontinuation	5	2	3	1	1	1	2	23
after dose reduction	0	0	0	1	0	0	2	18
without dose change	2	1	2	0	0	0	5	32
>2 to ≤3x ULN n (%)	12 (6.6)	0	12 (6.6)	2 (1.5)	0	7 (5.3)	22 (4.4)	74 (14.9)
Reversed to ≤2x ULN	12	-	11	2	-	6	20	70
Normalized to ≤ 1.2x ULN	10	-	9	2	-	6	19	61
after discontinuation	2	-	4	0	-	2	3	8
after dose reduction	0	-	0	0	-	2	1	12
without dose change	8	-	5	2	-	2	15	41
>1.2 to ≤2x ULN n (%)	32 (17.6)	8 (6.8)	20 (11.0)	23 (17.3)	10 (10.9)	14 (10.5)	72 (14.4)	84 (16.9)
Reversed to ≤2x ULN	n.a.[4]	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Normalized to ≤1.2x ULN	31	7	16	22	9	13	71	79
after discontinuation	6	3	0	4	1	2	5	10
after dose reduction	0	0	1	0	0	0	5	4
without dose change	25	4	15	18	8	11	61	65

[1] Only 10% of patients in MN302 received folate. All patients in US301 received folate.
[2] Includes MN303 extending the data to 12 months for the active treatment arms.
[3] The one leflunomide-treated subject in MN302 with an elevation >3x ULN that had not reversed to <2x ULN by the end of the study subsequently reversed on followup.
[4] n.a. = not applicable

In year-2, both ALT elevations and abnormal LFTs reported as adverse events occurred with lower frequency compared to year-1 as shown in Table 9. Two patients had ALT elevations >3x ULN that had not reversed to <2x ULN at the end of the study, but they subsequently reversed on followup.

Table 7. Clinically significant ALT elevation in year-1 and year-2: Phase III studies

	% of patients		
	LEF ITT cohort (N=824)	LEF Year-2 cohort (N=450)	
ALT			
>2 to ≤3 x ULN	4.6	5.1	2.9
>3 x ULN	3.0	2.4	1.8
Abnormal LFTs reported as AEs	7.8	5.6	3.3

ULN = upper limit of normal range, NA = not applicable

In the year 2 cohort a subject with an elevation in year 1 and year 2 is counted twice.

Table 8. US4001 Liver Enzyme Elevations in Combination Therapy with Methotrexate: Month 0-6 placebo-controlled study					
ALT (SGPT)	**>1.2 to ≤2x ULN n (%)**	**>2 to ≤3x ULN n (%)**	**>3x ULN n (%)**	**Total >1.2x ULN n (%)**	**Total >2x ULN n (%)**
LEF+MTX (N=130)	28 (21.5)	8 (6.2)	5 (3.8)	41 (31.5)	13 (10.0)
PLA+MTX (N=133)	6 (4.5)	2 (1.5)	1 (0.8)	9 (6.8)	3 (2.3)
AST (SGOT)					
LEF+MTX (N=130)	16 (12.3)	4 (3.1)	2 (1.5)	22 (16.9)	6 (4.6)
PLA+MTX (N=133)	5 (3.8)	0 (0.0)	1 (0.8)	6 (4.5)	1 (0.8)

C

APPENDIX C
LYMPHOMA/RA CITATIONS

Arellano F & Krupp P Malignancies in rheumatoid arthritis patients treated with cyclosporin *A Br J Rheumatol* 1993;32 Suppl:72-5

Baker GL, et al. Malignancy following treatment of RA with cyclophosphamide. *Am J Med* 1987; 83:1-9

Baltus J, et al: Occurrence of malignancies in patients with RA receiving cyclophosphamide. *Ann Rheum Dis* 1983; 42:368-73

Bleyer WA Editorial: Methotrexate induced Lymphoma? *J Rheumatol* 1997; 25:404-407

Cash JM & Klippel. JH Rheumatic Disease and cancer. *Clin Exp Rheumatol* 1991; 9:109-12

Casoli P & Tumiati B: Acute myelogenous leukemia in a rheumatoid arthritis patient under cyclosporine A therapy [letter]. *Clin Rheumatol* 1994;13:646-7

Cibere J, et al. Rheumatoid arthritis and risk of malignancy. *Arthritis Rheum* 1997; 40: 1580-86

Cleary ML, et al. Clonal analysis of transplant-associated lymphoproliferations based on the structure of the genomic termini of the Epstein-Barr virus. *Blood* 1988; 72:349

Cobeta-Garcia JC et al. Non-Hodgkin's lymphoma, rheumatoid arthritis and methotrexate [letter; comment]. *J Rheumatol* 1993;20:200-2

Cockburn I. Assessment of the risks of malignancy and lymphomas developing in patients using sandimmune. *Transplant Proceedings* 1987; 19: 1804-7

Cockburn ITR & Krupp P. The risk of neoplasms in patients treated with cyclosporine A. *J Autoimmun* 1991; 2:723-31

Davies JMS. et al: Lymphomatous changes during methotrexate therapy. *Arthritis Rheum* 1995; 38:S204

del Rincón I, et al. High Incidence of cardiovascular events in a rheumatoid arthritis cohort not explained by traditional cardiac risk factors. *Arthritis Rheum* 2001;44(12):2737-2745

Ellman MH et al. Lymphoma developing in a patient with RA taking low dose weekly methotrexate. *J Rheumatol* 1991; 18:1741-3

Emery R & Lake K. Post-transplantation lymphoproliferative disorder and OKT3. *NEJM* 1991; 324:1437-1439

Ferraccioli G.et al., Anticardiolipreschibing information antibodies in rheumatoid patients treated with etanercept or conventional combination therapy: direct and indirect evidence for a possible association with infections, *Annals of Rheumatic Disease;* 2002 61(4): 358-61

Ferraccioli GF et al. Epstein Barr virus associated Hodgkins lymphoma in a rheumatoid arthritis patient treated with methotrexate and cyclosporin A. *Arthritis Rheum* 1995; 38: 867-8

Frizzera G. Immunosuppression, autoimmunity and lymphoproliferative disorders. *Human Path* 1994; 25:627-9

Georgescu L, et al. Lymphoma inpatients with rheumatoid arthritis: association with disease state or methotrexate treatments. *Semin Arth Rheum* 1997; 26: 794-804

Gridley G, et al. Incidence of cancer among men with the Felty Syndrome. *Ann Intern Med* 1994; 120:35-39

Gridley G, et al. Incidence of cancer among patients with rheumatoid arthritis. *J Natl Cancer Inst* 1993; 85:307-11

Hakulinen T, et al. RA and cancer studies based on linking nationwide registries in Finland. *Am J Med* 1985; 78 (suppl 1A): 29-32

Hazleman BL. The comparative incidence of malignant disease in rheumatoid arthritics exposed to different treatment regimens. *Ann Rheum Dis* 1982; 41 (supp 12) :12-17

Hickey A, Urowitz M, Members of RAAR Steering Committee. The Rheumatoid Arthritis Azathioprine Registry (RAAR) - Interim Analysis of Malignancy and Mortality. *Arthritis Rheum* 1996; 39:S281

Isomaki HI, et al. Excess risk of lymphomas, leukemia and myeloma in patients with RA. *J Chron Dis* 1978; 31: 691-8

Jones M, et al. Does exposure to immunosuppressive therapy increase the 10 year malignancy and mortality risks in rheumatoid arthritis? *Br J Rheumatol* 1996; 35:738-45

Kahn MF, et al. Acute leukemia after treatment using cytotoxic agents for rheumatologic purposes. 19 cases among 2006 patients. *Nouv Presse Med* 1979; 8:1393-7

Kamel OW, et al. A population based case control study of non Hodgkin's lymphoma in patients with rheumatoid arthritis. *J Rheumatol* 1999; 26(8):1676-80

Kamel OW, et al. Brief report: reversible lymphomas associated with EBV occurring during methotrexate therapy for RA and dermatomyositis. *NEJM* 1993; 328:1317-1321

Kamel OW, et al. Immunosuppression-associated lymphoproliferative disorders in rheumatic patients. *Leuk Lymphoma* 1995;16:363-8

Kamel OW, et al. Lymphoid neoplasms in patients with rheumatoid arthritis and dermatomyositis: frequency of Epstein-Barr virus and other features associated with immunosuppression. *Hum Pathol* 1994;25:638-43

Kamel OW, et al. Hodgkin's Disease and Lymphoproliferations resembling Hodgkin's Disease in patients receiving long term low dose methotrexate therapy. *Am J Surg Path* 1996; 20: 1279-87

Kamel OW. Editorial: Lymphomas during long term methotrexate therapy. *Arch Dermatol* 1997; 133: 903-4

Kamel OW. Iatrogenic lymphoproliferative disorders in nontransplantation settings. *Sem Diagnostic Path* 1997; 14: 27-34

Katusic S, et al. Occurrence of malignant neoplasms in the Rochester Minnesota RA cohort. *Am J Med* 1985; 78 (suppl 1A): 50-51

Kingsmore SF, et al. Association of methotrexate, rheumatoid arthritis and lymphoma: report of two cases and literature review. *J Rheumatol* 1992; 19:1462-5

Kinlan LJ. Incidence of cancer in RA and other disorders after immunosuppressive treatment. *Am J Med* 1985; 78 (suppl 1A): 44-49

Kinlen LJ. Malignancy in autoimmune diseases. *J Autoimun* 1992; 5 (Suppl A):363-371

Kremer J. Editorial: Is methotrexate oncogenic in patients with rheumatoid arthritis? *Semin Arth Rheum* 1997; 26: 785-7

Laakso M , et al. Cancer Mortality in patients with RA. *J Rheumatol* 1986; 13:522-6

Lewis P, et al. Cause of death in patients with RA with particular reference to azathioprine. *Ann Rheum Dis* 1980; 39:457-61

Liote F, et al. Methotrexate related B lymphoproliferative disease in a patient with rheumatoid arthritis. Role of Epstein-Barr virus infection. *J Rheumatol* 1995; 22:1174-8

Moder KG, et al. Hematologic malignancies and the use of methotrexate in rheumatoid arthritis: a retrospective study. *Am J Med* 1995; 99:276-281

Mohan N, et al: Demyelination occurring during anti TNFa therapy for inflammatory arthritides. *Arthritis Rheum* 2001; 44: 2862-9

Morris CR & Morris AJ. Localized lymphoma in a patient with rheumatoid arthritis treated with parenteral methotrexate [letter]. *J Rheumatol* 1993; 20:2172-3

Myllykangas-Luosujarvi R, et al. Death attributed to antirheumatic medication in a nationwide series of 1666 patients with rheumatoid arthritis who have died. *J Rheumatol* 1995; 22:2214-7

Natkunam Y, et al. Epstein Barr virus strain type and latent membrane protein 1 gene deletions in lymphomas in patients with rheumatic diseases. *Arthritis Rheum* 1997; in press

Paul C, et al. Epstein Barr virus associated lymphoproliferative disease during methotrexate therapy for psoriasis. *Arch Dermatol* 1997; 133: 867-71

Penn I. Cancers complicating organ transplantation. *NEJM* 1990; 323: 1767-176

Pointud P, et al. Acute leukemia after low dose methotrexate therapy in a patient with rheumatoid arthritis. *J Rheumatol* 1993; 20:1215-6

Porter D, et al. Non-Hodgkins lymphoma in rheumatoid arthritis. *Ann Rheum Dis* 1991; 50:275-6

Prior P. Cancer and RA: Epidemiologic Considerations. *Am J Med* 1985; 78 (suppl 1A):15-21

Radis CD, et al. Effects of cyclophosphamide on the development of malignancy and on long term survival of patients with RA. A 20 year follow up study. *Arthritis Rheum* 1995; 38:1120-7

Shiroky JF, Frost A, Skelton JD et al: Complications of immunosuppression associated with weekly low dose methotrexate. J Rheumatol 1991; 18:1172-5

Shiroky JF & Newkirk MM. Reversible lymphomas (letter). *NEJM* 1993; 329: 1657-8

Silman AJ, et al. Lymphoproliferative cancer and other malignancy in RA patients treated with azathioprine: a 20 year follow up study. *Ann Rheum Dis* 1988; 47:988-92

Singh G, et al. Toxic effects of azathioprine in rheumatoid arthritis: a national post-marketing perspective. *Arthritis Rheum* 1989; 32:837-843

Symmons D. Excess mortality in rheumatoid arthritis -- is it the disease or the drugs? *J Rheumatol* 1995; 22:2200-2

Symmons D. Neoplasia in RA. *J Rheumatol* 1988; 15:1319-22

Symmons D. Neoplasms of the immune system in RA. *Am J Med* 1985; 78 (suppl 1A): 22-28

Taillan B, et al. Lymphoma developing in a patient with rheumatoid arthritis taking methotrexate. *Clin Rheumatol* 1993; 12:93-4

Tennis P, et al. Record linkage to conduct an epidemiologic study on the association of RA and lymphoma in the province of Sasketchewan Canada. *J Clin Epidemiol* 1993; 46: 685-95

Thoburn R & Katz P, coeditors, Hotline; American College of Rheumatology, June 1, 1995

Thomason RW, et al. Epstein-Barr virus and lymphoproliferation in methotrexate-treated rheumatoid arthritis. *Mod Pathol* 1996; 9:261-6

van de Rijn M, et al. Epstein-Barr virus clonality in lymphomas occurring in patients with rheumatoid arthritis. *Arthritis Rheum* 1996;39:638-42

Van den Borne BEEM, et al. No increased risk of malignancies and mortality in cyclosporin A-treated patients with rheumatoid arthritis. *Arthritis Rheum* 1998; 41:1930-7

Weinblatt, ME. The experience with CAMPATH 1H. Pp 161-8 In: Strand V, Simon L, Pillemer S, ed. Biologic Agents In Autoimmune Diseases IV. Arthritis Foundation, 1996

Williams CA, et al. Lymphoma and leukemia in RA: a matched case control study in the ARAMIS population. *Arthritis Rheum* 1995; 38:S204

Zijlmans JM, et al. Epstein-Barr virus-associated lymphoma in a patient with rheumatoid arthritis treated with cyclosporine (letter). *NEJM* 1992; 326:1363

Zimmer-Galler I & Lie JT. Choroidal infiltrates as the initial manifestation of lymphoma in RA after treatment with low dose methotrexate. *Mayo Clin Proc* 1994; 69:258-61

STUDY SYNOPSIS
HWA 486/F/USA/301/RA
ULTRA Study (Utilization of Leflunomide for Treatment of Rheumatoid Arthritis)
2 Year and Alternate Therapy

Title
A Phase III, Double Blind, Randomized, Placebo-Controlled Study to Compare the Activity and Safety of Leflunomide to Methotrexate or Placebo in Subjects with Active Rheumatoid Arthritis (Final 24-Month Data Including Alternate Therapy)

Investigators, study sites
Multinational/47 centers: United States of America (42), Canada (5)

Study dates
May 30, 1995 – November, 1998

Report type
Clinical/biometric report, final. This is the second final study report for this protocol. The first study report included results from the first year of therapy, as this was the primary efficacy endpoint and was used to support applications submitted globally for approval of leflunomide. The current report presents data from two years of therapy (the total time frame of the protocol) and from the alternate therapy phase of the protocol).

Report origin
Quintiles, San Diego, CA, USA

Hoechst Marion Roussel Deutschland GmbH, Frankfurt, Germany

Date of issue
August 25, 1999

Phase III

Indication
Rheumatoid arthritis (RA)

Study objectives
Primary objectives: To compare the efficacy and safety of leflunomide (LEF) with placebo (PBO) in subjects with active RA who had never previously received methotrexate (MTX).

Secondary objectives: To compare the efficacy and safety of LEF with MTX in subjects with active RA.

Study design
Phase III, multinational, multicenter, double-blind, randomized, placebo- and MTX-controlled study of parallel group design

Study medication and dosage

	LEF (once per day)	PBO	MTX (once per week)
Drug formulation/ Frequency	100 mg, Days 1-3 20 mg, Days 4-7 20 mg, Wks 2-6 20 mg if tolerated*, Wks 7-104	Matching placebo dispensed at times scheduled for LEF or MTX	7.5 mg, Wks 1-6 7.5 or 10 mg**, Wk 7 7.5 or 12.5 mg**, Wk 8 7.5 or 15 mg**, Wks 9-52 7.5, 15, 17.5 or 20 mg, Wks 53-104

* Dose could be lowered to 10 mg/day depending on tolerance.
** If active disease was present at Week 6.
All subjects were to receive folate (1 mg, twice daily).

Duration of treatment
Up to 104 weeks

Study population
Subjects with a diagnosis of RA by ACR criteria of ≥ 6 months' duration, who had active RA by ACR criteria at screening and baseline

Study variables

Efficacy
Primary. ACR20 responder rate: Percentage of subjects who met the ACR20 responder criteria at endpoint.

Secondary. Tender joint count, swollen joint count, patient global assessment, physician global assessment, modified Health Assessment Questionnaire (MHAQ), pain intensity assessment, erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), morning stiffness, X-rays of hands and feet, questionnaires on quality of life and functional ability (Medical Outcomes Study (MOS) SF-36, MOS current health perceptions scale, Work Limitations Questionnaire, standard HAQ, and Problem Elicitation Technique (PET))

Safety
Adverse events; hematology, blood chemistry, urinalysis; physical examination, supine blood pressure, heart rate, oral body temperature, body weight, 12-lead ECG, and chest X-ray.

Pharmacokinetics
Blood samples for determination of plasma concentrations of A77 1726 and TFMA. Results will be presented in a separate report.

Pharmacoeconomics
Utilization of health-care resources; studied for first 52 weeks only. Results were presented in a previous report.

Statistical methods
The statistical analyses were performed for three cohorts of subjects:
(a) intent-to-treat cohort (ITT cohort), (b) the subset of subjects with data during year 2 of therapy

(year-2 cohort), and (c) subjects enrolled in the alternate therapy phase. In addition to descriptive statistics, the following tests were performed:

Efficacy variables

Primary: ACR20 responder rates at endpoint in the ITT cohort analyzed by logistic regression.

Secondary: All secondary efficacy variables: analysis of covariance (ANCOVA) for all those cohorts analyzed.

Results

Study sample

In total, 511 subjects were randomized; 3 were randomized but not treated, 508 were randomized and treated (190 LEF, 128 PBO, and 190 MTX); 235 completed greater than 52 weeks of therapy (98 LEF, 36 PBO, and 101 MTX); 116 entered the alternate therapy 1-year treatment group (35 MTX to LEF, 56 PBO to LEF, and 25 LEF to MTX).

Of the treated subjects in the ITT cohort, 371 (73.0%) were females and 137 (27.0%) were males: LEF 72.6% females; PBO 71.9% females; MTX 74.2% females. Mean ages (years) in the 3 treatment groups were similar: LEF 54.0±12.1, PBO 54.7±10.6, and MTX 53.3±11.6. The treatment groups were similar for duration of RA, age at onset of RA, and prior disease modifying antirheumatic drug (DMARD) use.

Of the treated subjects in the year-2 cohort, 162 (68.9%) were females and 73 (31.1%) were males: LEF 69.4% females, PBO 69.4% females, and MTX 68.3% females. Mean ages (years) in the 3 treatment groups were similar: LEF 55.2±11.7, PBO 54.2±11.4, and MTX 53.3±12.4. The treatment groups were similar for duration of RA, age at onset of RA, and prior DMARD use.

Of the treated subjects in the alternate therapy group, 90 (77.6%) were females and 26 (22.4%) were males: PBO/LEF 75.0% females, MTX/LEF 88.6% females, LEF/MTX 68.0% females. Mean ages (years) in the 3 treatment groups were similar: PBO/ LEF 56.7±11.1, MTX/LEF 52.4±10.8, and LEF/MTX 52.2±12.5.

Study regimen

Reasons for early withdrawal from the ITT cohort

Reason	LEF (N=190)		PBO (N=128)		MTX (N=190)		Total (N=508)
	N	%	N	%	N	%	
Lack of efficacy	37	19.5	71	55.5	54	28.4	162
Adverse event	52	27.4	12	9.4	30	15.8	94
Lost to follow-up	1	0.5	1	0.8	4	2.1	6
Protocol violation	2	1.1	2	1.6	1	0.5	5
Noncompliance	–	–	1	0.8	2	1.1	3
Death	–	–	–	–	2	1.1	2
Other	15	7.9	14	10.9	17	8.9	46
Total	107	56.3	101	78.9	110	57.9	318

Reasons for early withdrawal from the year-2 cohort

Reason	LEF (N=98)		PBO (N=36)		MTX (N=101)		Total (N=235)
	N	%	N	%	N	%	
Lack of efficacy	4	4.1	2	5.6	5	5.0	11
Adverse event	8	8.2	–	–	8	7.9	16
Lost to follow-up	–	–	1	2.8	2	2.0	3
Protocol violation	1	1.0	1	2.8	–	–	2
Noncompliance	–	–	–	–	1	1.0	1
Death	–	–	–	–	1	1.0	1
Other	2	2.0	5	13.9	4	4.0	11
Total	15	15.3	9	25.0	21	20.8	45

Efficacy

506 subjects were included in the ITT cohort (LEF 190, PBO 128, MTX 188); 235 subjects were included in the year-2 cohort (LEF 98, placebo 36, MTX 101).

LEF vs. placebo.

In the ITT cohort, ACR20 response rates at Month 24 (by LOCF) showed leflunomide to be statistically significantly better than placebo. Analyses of mean changes in the components of the ACR20 response over time showed leflunomide to be highly statistically superior to placebo for all measures.

Month 24 ACR20 response rates (LOCF): ITT Cohort			
Treatment	No. of Subjects ACR20 Responders/Total	ACR20 Response Rate	p-value (95% confidence interval)
LEF	99/186	53.2	LEF vs PL: p≤0.001, 95% CI 21.7% to 41.9
PBO	34/128	26.6	LEF vs MTX: p=0.317, 95% CI-5.0% to 15.3%
MTX	90/188	47.9	MTX vs PL: p≤0.001, 95% CI 16.7% to 36.7%

In the year-2 cohort, ACR20 response rates were high in all groups since the year-2 cohort was enriched by subjects with positive treatment effect continuing into a second year of therapy. The size of the placebo group (36, 28% of the originally enrolled subjects) prohibited statistical comparisons with that group. Analyses of mean changes in the components of the ACR20 response over time showed similarly high improvement in the signs and symptoms of RA in all treatment groups, with the greatest relief seen in the leflunomide group. The year-2 cohort results indicate sustained effect over two years of treatment.

12 and 24 Month ACR20 response rates (LOCF): Year-2 Cohort					
Treatment	12 Month No. of ACR20 Responders/Total	ACR20 Response Rate	24 Month No. of ACR20 Responders/Total	ACR20 Response Rate	p-value (95% confidence interval)
LEF	75/97	77.3	77/97	79.4	LEF vs MTX:
PBO	22/36	61.1	23/36	63.9	p=0.049, 95% CI
MTX	61/101	60.4	68/101	67.3	0.1% to 24.4%

Statistical methods: Logistic regression: 95% CI for LEF vs MTX.

Analyses of the Sharp scores for X-rays of hands and feet for the ITT cohort at 12 months showed that increase in the total Sharp score was statistically significantly lower in the leflunomide group than the placebo group (0.5 vs 1.9, p=0.0016) This was also demonstrated in joint-space-narrowing subscore; no difference for the mean change in the erosion subscore was found. For the year-2 cohort, the Sharp scores showed that in subjects who continued into year 2, little progression was seen in all treatment groups at both year 1 and year 2, indicating continued protection against joint deterioration.

Analyses of the functional ability and health-related quality of life measures showed that the use of LEF for up to 104 weeks maintained improvements in physical function and health-related quality of life demonstrated after one year of treatment.

LEF vs. MTX.

In the ITT cohort, the ACR20 response rates for the LEF and MTX groups were statistically equivalent. The ACR20 response rate for LEF was 53.2% and for MTX was 47.9%. Response rates over time showed the LEF response occurred earlier and was higher than in either of the other two treatment groups. The MTX response was statistically significantly better than the placebo response. In the year-2 cohort, the ACR20 response rates for the LEF and MTX groups were 79.4% and 67.3% respectively, which were not statistically equivalent.

The X-ray analysis of hands and feet for the ITT cohort showed LEF and MTX to be statistically equivalent. The mean increase in total Sharp score was statistically significantly lower in the MTX group than in the placebo group. For the year-2 cohort, the results showed similar and low increases in total Sharp scores in both groups over year 2.

Functional ability and health-related quality of life measures showed better responses in the LEF group than in the MTX group on multivariate analysis, demonstrating overall statistical significance.

Safety

Adverse events.

In the ITT cohort, there were 84 (16.5%) serious adverse events; they were reported more frequently in the LEF (18.9%) and MTX (18.9%) than in the PBO (9.4%) groups. Few of these events were considered related to study drug (LEF 1.6%, PBO 1.6, and MTX 3.7%). In the year-2 cohort, there were 53 (22.6%) serious adverse events; they were reported more frequently in the MTX (24.8%) than the LEF (22.4%) and PBO (16.7%) groups. Few of these events were considered related to study drug (LEF 1.0%, MTX 2.0%). There were no deaths on study in the LEF, 2 in the MTX (1 study drug related), and 1 in the PBO groups.

Withdrawals due to serious adverse events were lower in the LEF compared to the MTX group in both the ITT and year-2 cohorts (LEF 4.2%, PBO 1.6%, and MTX 6.3%), and (LEF 1.0%, PBO 0%, and MTX 5.0%), respectively. Serious adverse events which occurred in > 1% LEF subjects in the ITT cohort were pneumonia (4 subjects), infection (2), joint disorder (5), cholelithiasis (4), hypertension (2), and deep thrombophlebitis (2); for the year-2 cohort, they were cholelithiasis (4), joint disorder (4), infection (2), hypertension (2), cholecystitis (2), and deep thrombophlebitis (2).

Adverse events (i.e., serious and non-serious events) in the ITT cohort were more frequent in the LEF (98.4%) than in the PBO (88.3%) or MTX (92.1%) groups. In the year-2 cohort, the frequency of adverse events was similar for all treatment groups: LEF (100.0%), PBO (94.4%), and MTX (96.0%). The most frequently reported adverse events, regardless of causality, in the

LEF group in the ITT cohort, were: respiratory infection (37.4%), diarrhea (36.8%), headache (20.0%), hypertension (18.4%), and rash (17.4%). In the PBO group: respiratory infection (25.0%), nausea (18.8%), headache (17.2%), and dyspepsia (16.4%). In the MTX group: respiratory infection (38.4%), headache (23.2%), diarrhea (21.6%), and nausea (20.5%). In the year-2 cohort, the most frequently reported adverse events, regardless of causality, in the LEF group were: respiratory infection (55.1%), diarrhea (43.9%), hypertension (28.6%), dyspepsia (24.5%), and rash (21.4%). In the PBO group: respiratory infection (38.9%), diarrhea (30.6%), dyspepsia (27.8%), and accidental injury (25.0%). In the MTX group: respiratory infection (54.5%), diarrhea (23.8%), headache (23.8%), and accidental injury (20.8%).

In the ITT cohort, adverse events considered related to study drug occurred in LEF (80.5%), PBO (56.3%), and MTX (65.8%) of subjects. The frequency of withdrawals due to adverse events was higher in the LEF (26.8%) than in the PBO (9.4%) or MTX (16.8%) groups. For the year-2 cohort, adverse events considered related to study drug occurred more frequently in the LEF (81.6%), compared to the PBO (63.9%), and MTX (62.4%) groups. The frequency of withdrawals due to adverse events was similar in the LEF (7.1%) compared to the MTX (8.9%) group, with PBO (0%).

In the ITT cohort, the most common adverse events considered related to LEF administration were of gastrointestinal origin, predominantly diarrhea (27.9% LEF, 13.3% PBO, and 13.7% MTX), LFT abnormalities (15.3% LEF, and 10.5% MTX), dyspepsia (LEF 13.2%, PBO 13.3%, and MTX 8.9%) and nausea (LEF 12.6%, PBO 18.0%, and MTX 15.8%). In the year-2 cohort, the most common adverse events considered related were diarrhea (LEF 31.6%, PBO 22.2%, and MTX 11.9%), dyspepsia (LEF 18.4%, PBO 16.7%, and MTX 8.9%), rash (LEF 13.3%, PBO 2.8%, and MTX 3.0%), alopecia (LEF 13.3%, PBO 0%, and MTX 4.0%), hypertension (LEF 12.2%, PBO 5.6%, and MTX 1.0%), abdominal pain, digestive system (LEF 10.2%, PBO 5.6%, and MTX 7.9%), and LFT abnormalities (LEF 10.2%, PBO 2.8%, and MTX 8.9%).

Of the 63 total subjects in the ITT cohort with adverse events leading to withdrawal from the study which were related, 38 (20.0%) received LEF, 7 (5.5%) PBO, and 18 (9.5%) MTX. Of these, the most frequent events in the LEF group (21 subjects, 11.1%) were in the digestive system, and included LFT abnormalities, diarrhea, and nausea. Of the 6 total subjects in the year-2 cohort with adverse events leading to withdrawal from the study which were related, 4 (4.1%) received LEF, 0% PBO, and 2 (2.0%) MTX.

A comparison of the incidence of adverse events in year 1 with year 2 in the year-2 cohort showed that, in general, adverse events decreased in year 2 in all treatment groups. Exceptions were hypertension increased in LEF from 15.3% to 23.5%, PBO from 5.6% to 11.1%, and MTX from 3.0% to 4.0%; arthralgia increased in LEF from 5.1% to 11.2%; and peripheral edema increased in LEF from 5.1% to 10.2%.

The highest incidence of grouped adverse events in the leflunomide treatment group in the ITT cohort was reported in the gastrointestinal system. Diarrhea was the most frequently reported: LEF (36.8%), PBO (20.3%), and MTX (21.6%). The majority of diarrhea adverse events in the LEF group were mild to moderate, all resolved without sequelae, and occurred during the first 1-2 months of study drug administration, an effect that may have been related to the loading dose of leflunomide at the initiation of study. Study treatment was decreased in 5 (2.6%) LEF subjects; interrupted in 3 (1.6%) LEF, 1 (0.8%) PBO, and 6 (3.2%) MTX subjects; and discontinued in 18 (9.5%) LEF, 3 (2.3%) PBO, and 10 (5.3%) MTX subjects. In the year-2 cohort, diarrhea was also the most frequently reported: LEF (43.9%), PBO (30.6%), and MTX (23.8%). These adverse events in the LEF group were mild to moderate, and, with the exception of 2 (2%) subjects, all

resolved. Study treatment was decreased in 4 (4.1%) LEF, and 1 (1.0%) MTX subjects; interrupted in 2 (2.0%) LEF and 2 (2.0%) MTX subjects; and discontinued in 6 (6.1%) LEF and 3 (3.0%) MTX subjects. Diarrhea appeared to be an adverse effect of LEF, however, in both cohorts the majority of the diarrhea was mild to moderate.

In the ITT cohort, nausea/vomiting was reported at a similar rate for the three treatment groups: LEF (23.7%), PBO (22.7%), and MTX (21.6%). The majority of cases were mild to moderate, and 5 (2.6%) LEF and 5 (2.6%) MTX subjects had severe cases. Study treatment was decreased in 1 (0.5%) LEF and 1 (0.5%) MTX subjects; interrupted in 4 (2.1%) LEF, 1 (0.8%) PBO, and 6 (3.2%) MTX subjects; and discontinued in 19 (10.0%) LEF, 2 (1.6%) PBO, and 10 (5.3%) MTX subjects. In the year-2 cohort, nausea/vomiting was reported at a slightly lower rate in the LEF (18.4%) compared to PBO (25.0%) and MTX (20.8%) groups. Most cases were mild to moderate in severity and 2 (2.0%) LEF, and 1 (2.8%) PBO subjects required treatment. Study treatment was decreased in 1 (1.0%) LEF and 1 (1.0%) MTX subjects; interrupted in 1 (1.0%) MTX subject; and discontinued in 3 (3.1%) LEF and 4 (4.0%) MTX subjects. The incidence of nausea/vomiting and dyspepsia in both the ITT and year-2 cohorts were similar among treatment groups and may have been associated with NSAID administration. Nausea/vomiting did not appear to be an adverse effect of leflunomide administration.

Abdominal pain in the ITT cohort occurred at a slightly higher rate in LEF (18.9%) compared to PBO (10.9%) and MTX (14.7%) groups, and in the year-2 cohort, LEF (23.5%), PBO (22.2%), and MTX (16.8%). The similar occurrence of abdominal pain in all treatment groups in both cohorts may reflect NSAID use. Abdominal pain did not appear to be an adverse effect of leflunomide administration in both cohorts.

Oral ulcerations are an expected adverse event with methotrexate administration and occurred less frequently in the LEF group compared to the MTX group in the ITT cohort: LEF (6.8%), PBO (5.5%), and MTX (10.5%). In the year-2 cohort, the frequency was LEF (9.2%), PBO (2.8%), and MTX (14.9%).

In the ITT cohort, infections accounted for the second highest incidence of adverse events in all three treatment groups. Their occurrence was similar in the LEF (64.2%) and MTX (65.8%) groups, compared to the PBO (51.6%) group. The infections in the LEF group were generally mild to moderate. Treatment-related infections occurred in LEF (3.2%), MTX (2.1%), and PBO (0.8%) subjects. Study treatment was decreased in 5 (2.6%) LEF, and 6 (3.2%) MTX; interrupted in 6 (3.2%) LEF and 3 (1.6%) MTX subjects; and discontinued in 24 (12.6%) LEF, 5 (3.9%) PBO, and 18 (9.5%) MTX subjects. The most common infections were respiratory and their occurrence was similar in the LEF and MTX groups (37.4% and 38.4%, respectively), compared to PBO (25.0%) group. Respiratory infections related to study treatment were LEF (3.2%), MTX (2.1%), and PBO (0.8%). In the year-2 cohort, infections accounted for the highest incidence of grouped adverse events in all three treatment groups. Their occurrence was similar in the LEF (84.7%) and MTX (86.1%) groups, compared to the PBO (69.4%) group. Study treatment was decreased in 4 (4.1%) LEF and 6 (5.9%) MTX subjects; interrupted in 5 (5.1%) LEF and 3 (3.0%) MTX subjects; and discontinued in 7 (7.1%) LEF and 8 (7.9%) MTX subjects. The most common infections were respiratory and their occurrence was similar in the LEF and MTX groups (55.1% and 54.5%, respectively), compared to PBO (38.9%). Respiratory infections related to study treatment were LEF (5.1%), MTX (4.0%), and PBO (0%). There were no opportunistic infections or disseminated herpes zoster or herpes simplex.

Adverse events associated with the cardiovascular system in the ITT cohort were more frequently reported in the LEF (27.9%), compared to PBO (15.6%) and MTX (11.1%) groups. The most

frequently reported adverse events were hypertension (LEF 18.4%, PBO 8.6%, and MTX 4.7%), and chest pain (LEF 7.9%, PBO 6.3%, and MTX 4.7%). The majority of these events in the LEF group were mild to moderate. Hypertension adverse events were related to study drug in 8.9% LEF, 4.7% PBO, and 0.5% MTX subjects. There were 2 (1.1%) serious adverse events associated with hypertension in the LEF, 1 (0.8%) PBO, and none in the MTX groups. Hypertension at baseline was reported at a higher frequency in the LEF subjects (13.7%), compared to PBO (8.6%), and MTX (2.1%) subjects. New-onset hypertension was reported in 4.7% LEF, 0% PBO, and 2.6% MTX subjects. In the year-2 cohort, cardiovascular adverse events were more frequently reported in the LEF (38.8%), compared to PBO (25.0%) and MTX (9.9%) groups. The most frequently reported adverse events were hypertension (LEF 28.6%, PBO 13.9%, and MTX 5.9%), and chest pain (LEF 9.2%, PBO 11.1%, and MTX 4.0%). The majority of these events in the LEF group were mild to moderate. Hypertension adverse events were possibly or probably related to study drug in 12.2% LEF, 5.6% PBO, and 1.0% MTX subjects. There were 2 (2.0%) serious adverse events associated with hypertension in the LEF group. Hypertension at baseline was reported in 21.4% LEF, 13.9% PBO, but only 2.0% MTX subjects. New-onset hypertension was reported in 7.1% LEF, 0% PBO, and 4.0% MTX subjects. In both the ITT and year-2 cohorts, hypertension at baseline and concomitant NSAID and steroid use were higher in the LEF compared to the PBO and MTX groups, which may reflect the higher occurrence of hypertension in the LEF group. Leflunomide administration did not appear to have any clinically significant effect on blood pressure in the ITT and year-2 cohorts.

In the ITT cohort, potential allergic reactions were reported in 29.5% LEF, 16.4% PBO, and 22.1% MTX subjects. The most commonly reported adverse events were rash (17.4% LEF, 8.6% PBO, and 11.1% MTX) and allergic reactions (10.5% LEF, 4.7% PBO, and 6.3% MTX). Potential allergic reactions related to leflunomide administration consisted mainly of pruritus (3.7%) and rash (1.6%). Study treatment was decreased in 3 (1.6%) LEF and 2 (1.1%) MTX subjects; interrupted in 2 (1.1%) LEF, 1 (0.8%) PBO, and 2 (1.1%) MTX subjects. There were no serious adverse events. Treatment was discontinued for rash in 3 (1.6%) LEF, 3 (2.3%) PBO, and 0% MTX subjects. In the year-2 cohort, potential allergic reactions were reported in 38.8% LEF, 13.9% PBO, and 29.7% MTX subjects. The most commonly reported adverse events were rash (21.4% LEF, 8.3% PBO, and 12.9% MTX) and allergic reactions (17.3% LEF, 8.3% PBO, and 8.9% MTX). Study treatment was decreased in 3 (3.1%) LEF and 2 (2.0%) MTX subjects; interrupted in 1 (1.0%) LEF subject; and discontinued in 4 (4.1%) LEF and 1 (1.0%) MTX subjects. Potential allergic reactions related to leflunomide administration were higher than in the ITT cohort and consisted mainly of rash (13.3%), and pruritus (5.1%). No anaphylactic reactions or angioedema were noted in the ITT and year-2 cohorts.

In the ITT cohort, the incidence of rheumatoid arthritis (RA)-related adverse events were evenly distributed among the three treatment groups (26.3% LEF, 23.4% PBO, and 28.4% MTX). The majority of events were mild to moderate and were unrelated to study treatment. Vasculitis occurred in 1 (0.5%) LEF subject in year 1, which was unrelated to study drug administration, and 1 (0.5%) MTX subject in year 1, which was judged related to study drug administration. In the year-2 cohort, the incidence of RA-related adverse events was slightly higher than in the ITT cohort in the two active treatment groups (34.7% LEF and 38.6% MTX, compared to 25.0% PBO).

In the ITT cohort, the incidence of central nervous system adverse events was similar in all treatment groups (35.3% LEF, 28.1% PBO, and 35.3% MTX). Study treatment was decreased in 3 (1.6%) LEF, and 1 (0.5%) MTX subjects; interrupted in 1 (0.5%) LEF and 1 (0.5%) MTX subjects; and discontinued in 17 (8.9%) LEF and 17 (8.9%) MTX subjects. The most frequently reported event was headache (20.0% LEF, 17.2% PBO, and 23.2% MTX), of which related

events occurred in 12.1% LEF, 7.8% PBO, and 12.6% MTX. The occurrence of grouped adverse events of neuritis/neuropathy/paresthesia was similar in the two active treatment groups (7.9% LEF and 7.4% MTX). In the year-2 cohort, the incidence of central nervous system adverse events was slightly higher than in the ITT cohort, and was similar in all treatment groups (43.9% LEF, 38.9% PBO, and 45.5% MTX). Study treatment was decreased in 2 (2.0%) LEF and 1 (1.0%) MTX subjects; and discontinued in 5 (5.1%) LEF and 5 (5.0%) MTX subjects. The most frequently reported event was headache (19.4% LEF, 19.4% PBO, and 23.8% MTX), of which related events occurred in 10.2% LEF, 8.3% PBO, and 10.9% MTX. The occurrence of grouped adverse events of neuritis/neuropathy/paresthesia was similar in the two active treatment groups 13.3% LEF and 8.9% MTX).

Other adverse events of interest included alopecia, which occurred in the ITT cohort in 10.5% LEF, 0.8% PBO, and 5.8% MTX subjects. In the year-2 cohort, alopecia occurred in 13.3% LEF and 5.0% MTX subjects. Most cases were mild to moderate and resolved without treatment. There were 9 discontinuations from study treatment due to alopecia in the ITT cohort (3 LEF, 1 PBO, and 5 MTX subjects). In the year-2 cohort, there was only 1 discontinuation due to alopecia in the LEF and 1 decrease in dosage in the MTX group. It should be noted that the alopecia resolved in the LEF subjects who discontinued. Alopecia occurred with leflunomide and methotrexate administration, however the incidence in the LEF group declined from 12.2% in year 1 to 5.1% in year 2, compared to 5.0% and 3.0%, respectively for the MTX group.

Laboratory variables.

Leflunomide administration in both the ITT and year-2 cohorts did not appear to be associated with clinically significant changes in hemoglobin, hematocrit, RBC parameters, platelets, and WBC subpopulations. There was no clinically significant effect on sodium, potassium, chloride, bicarbonate, BUN or creatinine. Leflunomide administration did appear to be associated with a decrease in serum uric acid in both the ITT and year-2 cohorts, due to the known uricosuric effect of leflunomide on the brush border membrane of the proximal renal tubule cells. In both the ITT and year-2 cohorts, leflunomide did not appear to be associated with clinically significant changes in total protein, albumin, and total bilirubin.

Leflunomide administration appeared to be associated with elevations of SGPT (ALT) and SGOT (AST) in the ITT and year-2 cohorts. In the ITT cohort, there were 10 (5.3%) LEF subjects with SGOT (AST) > 2x ULN to 3x ULN, and all normalized to ≤ 1.2x ULN. Similarly, of 12 (6.3%) LEF subjects with SGPT (ALT) > 2x to 3x ULN, all reversed to ≤ 2x ULN, 11 (5.8%) normalized to ≤ 1.2x ULN, 3 (1.6%) normalized after discontinuation, and 8 (4.2%) normalized without dose reduction of study treatment. There were 7 (3.7%) LEF subjects with SGOT (AST) > 3x ULN, and all normalized to ≤ 1.2x ULN. Similarly, of 12 (6.3%) LEF subjects with SGPT (ALT) > 3x ULN, all reversed to ≤ 2x ULN, 11 (5.8%) normalized to ≤ 1.2x ULN, 6 (3.2%) normalized after discontinuation, and 5 (2.6%) normalized without dose reduction of study treatment. Liver function abnormalities that were judged related to study drug were similar in the two active treatment groups, and occurred in 30 (15.8%) LEF, 4 (3.1%) PBO, and 22 (11.6%) MTX subjects. Of these, 14 (7.4%) LEF, 1 (0.8%) PBO, and 7 (3.7%) MTX subjects discontinued from the study. In the year-2 cohort, SGPT (ALT) was moderately elevated at any visit in 8.2% LEF, 5.6% PBO, and 5.0% MTX subjects; and marked elevations occurred in 6.1% LEF, 5.6% PBO, and 4.0% MTX subjects. SGOT (AST) was moderately elevated in 6.1% LEF, 0% PBO, and 5.0% MTX subjects; marked elevations occurred in 3.1% LEF, 2.8% PBO, and 1.0% MTX subjects; and elevations > 8x ULN occurred in 1.0% LEF subjects. At worst evaluation, there were 5.1% LEF subjects with SGOT (AST) > 2x ULN to 3x ULN, and all normalized to ≤ 1.2x ULN. Similarly, of 5.1% LEF subjects with SGPT (ALT) > 2x ULN to 3x

ULN, all normalized without dose reduction. There were 4.1% LEF subjects with SGOT (AST) > 3x ULN, and all normalized to ≤ 1.2 x ULN. Similarly, of 6 (6.1%) LEF subjects with SGPT (ALT) > 3x ULN, all normalized to ≤ 1.2x ULN without dose reduction. Liver function abnormalities that were judged related to study drug were very similar in the two active treatment groups, and occurred in 10 (10.2%) LEF, 1 (2.8%) PBO, and 9 (8.9%) MTX subjects. Of these, 1 (1.0%) LEF subject discontinued from the study.

Leflunomide administration did not appear to be associated with any clinically significant adverse effect on alkaline phosphatase in the ITT and year-2 cohorts.

In both the ITT and year-2 cohorts, there were no adverse effects of leflunomide administration on other chemistry parameters: LDH, triglycerides, total cholesterol, calcium, phosphorous, glucose, and creatine kinase. In both the ITT and year-2 cohorts, there was essentially no difference between the three treatment groups in the tested parameters of urinalysis.

Clinical variables.
In both the ITT and year-2 cohorts, there were no clinically significant differences between the three treatment groups in ECG, chest X-ray or physical examination results. Leflunomide administration had no effect on body temperature, weight, or heart rate.

Alternate therapy

Cohort

Efficacy
114 subjects were evaluable for efficacy in the alternate therapy phase of the protocol (56 PBO/LEF, 34 MTX/LEF, 24 LEF/MTX). Results were variable over time, but indicated that improvement occurred in all three treatment groups, not just those subjects entering from the placebo arm of initial therapy. Up to half of the subjects not responding to LEF or MTX responded well to the other DMARD. The percentage of subjects who were ACR20 responders at endpoint was somewhat higher in the PBO/LEF (52.8%) group than in the groups switching from one DMARD to the other (MTX/LEF group [36.4%] and LEF/MTX [50.0%]). ACR50 rates were higher in the PBO/LEF group (23.8%) than in the other two groups (MTX/LEF 21.2% and LEF/MTX 16.7%).

Safety
Serious adverse events were reported more frequently in the LEF/MTX group (20.0%) than in PBO/LEF (16.1%) or MTX/LEF (11.4%) groups, with only a few considered as related to study drug administration (4.0% LEF/MTX, 3.6% PBO/LEF, and 0% MTX/LEF). Withdrawals due to serious adverse events were infrequent in all treatment groups (4.0% LEF/MTX, 3.6% PBO/LEF, and 0% MTX/LEF). Serious adverse events that occurred in more than one subject receiving leflunomide included coronary artery disorder (2 subjects), bone necrosis (2), and joint disorder (2). Adverse events (both serious and non-serious adverse events) were reported more frequently with leflunomide treatment (100.0% PBO/LEF, 94.3% MTX/LEF) than methotrexate (88.0% LEF/MTX).

Adverse events considered related to study drug administration were more frequent in the LEF groups (78.6% PBO/LEF, 77.1% MTX/LEF) than the methotrexate group (56.0% LEF/MTX). The frequency of withdrawals due to adverse events was higher in the LEF groups (19.6% PBO/LEF, 14.3% MTX/LEF) compared to methotrexate treatment (12.0% LEF/MTX), due to a

greater incidence of gastrointestinal disorders. The most common adverse events considered related to leflunomide administration were of gastrointestinal origin and consisted predominantly of diarrhea (26.8% PBO/LEF, 28.6% MTX/LEF, 8.0% LEF/MTX), nausea (16.1% PBO/LEF, 17.1% MTX/LEF, 16.0% LEF/MTX), dyspepsia (10.7% PBO/LEF, 8.6% MTX/LEF, 4.0% LEF/MTX), abdominal pain (8.9% PBO/LEF, 5.7% MTX/LEF, 8.0% LEF/MTX), and LFT abnormalities (7.1% PBO/LEF, 11.4% MTX/LEF, 8.0% LEF/MTX). Other adverse events that appeared related to leflunomide were alopecia (14.3% PBO/LEF, 5.7% MTX/LEF, 4.0% LEF/MTX) and rash (7.1% PBO/LEF, 8.6% MTX/LEF, 0% LEF/MTX).

Of the 15 total subjects with non-serious adverse events leading to withdrawal that were related to treatment, 16.1% were in the PBO/LEF, 11.4% in the MTX/LEF, and 8.0% in the LEF/MTX groups. Of these, the most frequent events (5 subjects, 8.9%) were in the digestive body system, and included LFT abnormalities, vomiting, diarrhea, nausea, and aphthous stomatitis.

The overall incidence of infections was higher in the methotrexate group compared to both LEF groups (64.0% LEF/MTX, 57.1% PBO/LEF, 54.3% MTX/LEF). The majority of infections as adverse events were respiratory infections, which occurred at a higher rate in the LEF/MTX group (44.0%), compared with the PBO/LEF (33.9%) and MTX/LEF (22.9%) groups.

The most frequently reported adverse events in the cardiovascular system were hypertension (14.3% MTX/LEF, 7.1% PBO/LEF, 0% LEF/MTX), and chest pain (11.4%, 5.4%, and 8.0%, respectively). Hypertension adverse events were related to study drug in 11.4% MTX/LEF and 1.8% PBO/LEF subjects. In the subset of subjects with hypertension as an adverse event, there was a higher incidence of concomitant hypertension at baseline in 2 (3.6%) PBO/LEF, and 4 (11.4%) MTX/LEF subjects, compared to 0% LEF/MTX subjects. New-onset hypertension was reported more frequently in LEF subjects (3.6% PBO/LEF, 2.9% MTX/LEF, 0% LEF/MTX), but the incidence was low. There were no treatment discontinuations due to hypertension.

Potential allergic reactions were reported in 28.6% PBO/LEF, 25.7% MTX/LEF, and 8.0% LEF/MTX subjects. Potential allergic reactions possibly or probably related to leflunomide consisted mainly of rash (7.1% PBO/LEF and 8.6% MTX/LEF, compared to 0% LEF/MTX).

The incidence of RA-related adverse events was evenly distributed among the three treatment groups (25.0% PBO/LEF, 22.9% MTX/LEF, and 24.0% LEF/MTX).

Central nervous system adverse events were reported in 32.1% PBO/LEF, 31.4% MTX/LEF, and 24.0% LEF/MTX subjects. The majority of central nervous system adverse events consisted of headache (21.4% PBO/LEF, 14.3% MTX/LEF, and 24.0% LEF/MTX), paresthesia (10.7%, 11.4%, and 4.0%), and dizziness (7.1%, 5.7%, and 0%). There was one serious adverse event in the MTX/LEF group, consisting of anxiety and paresthesia that was judged unrelated to study drug administration.

Laboratory variables

Leflunomide administration had no effect on sodium, potassium, chloride, bicarbonate, BUN or creatinine. Leflunomide appeared to be associated with a decrease in uric acid, due to the known effect of the drug on the brush border membrane of the proximal renal tubule cells without alterations in renal function or evidence of renal tubular acidosis.

All treatment groups at any visit had similar elevations of SGPT (ALT) and SGOT (AST). Most elevations were mild to moderate and resolved during treatment; moderate elevations of SGPT (ALT) (> 2x ULN to 3x ULN) were highest in the LEF/MTX group (8.0%), compared to the

PBO/LEF (3.6%), and MTX/LEF (5.7%) groups. Marked elevations at worst evaluation (>3x ULN) were less frequent and reversed without dose reduction. Liver function abnormalities that were judged related to study drug were similar in all treatment groups, and occurred in 4 (7.1%) PBO/LEF, 4 (11.4%) MTX/LEF, and 2 (8.0%) LEF/MTX subjects. Of these, 1 (1.8%) PBO/LEF, 1 (2.9%) MTX/LEF, and 2 (8.0%) LEF/MTX subjects discontinued from the study. Leflunomide administration did not appear to be associated with any clinically significant adverse effects on SGPT (ALT) and SGOT (AST) in the alternate therapy cohort.

Alkaline phosphatase was mildly elevated (> 1.2 x ULN and < 2 x ULN) in 8.9% PBO/LEF, 5.7% MTX/LEF, and 4.0% LEF/MTX subjects. Moderate elevation (> 2 x ULN to < 3 x ULN) occurred in 2.9% of MTX/LEF subjects. Leflunomide appeared to be associated with a mild, but clinically insignificant elevation of alkaline phosphatase in a small percentage of patients.

There was no effect on total protein, albumin or total bilirubin and no clinically significant differences in urinalysis results between the three treatment groups.

Leflunomide was associated with an increase in triglycerides, with shifts from normal values at baseline to values above the normal range at endpoint in 20.4% PBO/LEF, 26.9% MTX/LEF, and 0% LEF/MTX subjects. Leflunomide did not appear to be associated with any clinically significant increase in total cholesterol.

There were no adverse effects of leflunomide on other chemistry parameters, such as glucose, creatine kinase (CK), calcium, and phosphorous.

Clinical variables

There were no clinically significant differences between the three treatment groups in ECG, chest X-ray or physical examination results. Leflunomide had no effect on body temperature or heart rate. Clinically relevant changes in weight from baseline to endpoint occurred in 3.6% PBO/LEF, 8.6% MTX/LEF, and 0% of LEF/MTX subjects. Hence, weight loss in a small number of subjects may be associated with leflunomide.

Subjects in both the leflunomide treated groups in the alternate therapy phase had similar safety profiles, regardless of the treatment during the initial therapy phase.

Comments/Conclusions

This clinical trial yielded highly statistically significant results that demonstrated leflunomide was superior to placebo in the treatment of active RA. This was demonstrated by a reduction of signs and symptoms of RA, and the sustaining of clinical and radiographic benefit over 24 months. The retardation of disease progression was maintained between year 1 and year 2 in subjects who continued therapy. Analysis of the functional ability and health-related quality of life measures over 2 years showed the use of leflunomide effected statistically significant improvements over both short- and long-term therapy. In addition, ACR20 response rates for the leflunomide and methotrexate groups were statistically equivalent, though the leflunomide rate was numerically higher at 24 months. There was an earlier response in the LEF group compared to the MTX group, although the results for methotrexate approached those of leflunomide at endpoint. The methotrexate response was highly significantly better than the placebo response. The safety profile of leflunomide in the ITT cohort over 2 years, and the year-2, and alternate therapy cohorts appeared to be acceptable and was generally similar to that of methotrexate. These results provide a safety profile that supports the findings from the one year data.

Hoechst

STUDY SYNOPSIS
HWA 486/6/MN/305/RA

TITLE

Comparative extension trial of the safety and efficacy of leflunomide versus sulfasalazine in patients with active rheumatoid arthritis

INVESTIGATORS STUDY SITES

Multinational / 30 centers: Australia (1), Austria (3) Denmark (2) Germany (3) New Zealand (2) Norway (1) Slovenia (1) South Africa (4) The Netherlands (1) United Kingdom (12)

STUDY DATES

January 1995 - December 1997

REPORT TYPE

Clinical/biometric report, final

REPORT ORIGIN

ClinData International (Pty) Ltd, Bloemfontein, South Africa
Hoechst Marion Roussel Deutschland GmbH, Frankfurt, Germany

DATE OF ISSUE

3 November 1998

PHASE

III

INDICATION

Rheumatoid arthritis (RA)

STUDY OBJECTIVES

Primary objective: To investigate the safety of leflunomide (LEF) during long-term use in RA subjects.
Secondary objectives:

- To compare the efficacy and safety profiles of LEF with those of sulfasalazine (SSZ) after long-term treatment.
- To perform a pharmacoeconomic evaluation and investigate population pharmacokinetics.

STUDY MEDICATION AND DOSAGE

Drug	LEF	SSZ
Formulation	10 mg film-coated tablets	0.5 g enteric-coated tablets
Dosage	20 mg or 10 mg daily	2.0 or 1.5 g daily
Batch Nos.	See Informational Appendix B of study report	

As the LEF and SSZ treatments differed in appearance, placebos to both LEF and SSZ were used in a doubledummy technique to preserve the double-blind nature of the trial.

STUDY DESIGN

Multinational, parallel-group study with a double-blind treatment phase and a **12-week**, treatment-free observation phase. The study was an extension to the 24-week study HWA **486/6/MN/303/RA** ("study **303**") which in turn was an extension to the **24-week** study HWA **486/6/MN/301/RA** ("study 301"). The double-blind treatment period of study 305 continued until the last subject had completed 2 years of treatment in studies 301, 303, and 305. Subjects who received LEF or SSZ in study 301 continued on the respective medication in studies 303 and 305; subjects who received placebo in study 301 were switched to SSZ in a blinded manner at the start of study 303 and continued on SSZ in study 305 (**placebo/SSZ** group). At the start of study 305, all subjects continued on the same daily dosage of LEF and SSZ that they had been taking at completion of study 303.

STUDY POPULATION

Subjects with active RA who had completed 24 weeks of treatment in study 301 and 24 weeks of treatment in study 303 and who wished to continue treatment.

Study manager: Dr R Rosenburg **Biometricians: Ch Oed, A Grobler**
Authors: Dr R Rosenburg, Ch Oed, A Grobler **Medical writer: Dr G Schulz**

STUDY VARIABLES

Efficacy

X-ray evaluation of hands and **forefeet** (total Larsen X-ray score was the main variable), tender joint count, swollen joint count, investigator's global assessment of RA activity, patient's global assessment of RA activity, responder rates (**Paulus** and American College of Rheumatology [ACR]), treatment success rate, joint tenderness score, swollen joint score, duration of morning stiffness, pain intensity assessment, Health Assessment Questionnaire (HAQ), erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), rheumatoid factor (RF)

Safety

Observed and spontaneously reported adverse events; hematology, blood chemistry, urinalysis; physical examination, supine blood pressure, pulse, oral body temperature, body weight, la-lead electrocardiogram, chest X-ray

Pharmacokinetics

Trough blood samples for determination of plasma/serum levels of A77 1726.

Pharmacoeconomics

Utilization of health-care resources.
Results for both analyses to be presented in separate reports.

STATISTICAL METHODS

In addition to descriptive statistics, the following tests were performed:

Baseline comparability
Categorical variables: x^2 test (if cell frequency below 5, Fisher's exact test)
Continuous variables: analysis of variance (ANOVA)

Efficacy variables
X-ray evaluation of hands and forefeet (total Larsen score, subscores, number **of** eroded joints): analysis of covariance (ANCOVA). Responder and treatment success rates: logistic regression. All other variables: ANCOVA.
The main efficacy analysis was the intention-to-treat comparison of the mean changes in total Larsen X-ray score for hands and forefeet from baseline in study 301 to endpoint in study 305 between LEF and SSZ.

Safety variables
Time to withdrawal/Time to first occurrence of events: life-table analysis
Laboratory data: Wilcoxon signed rank test

RESULTS

Study sample

146 (87%) of the 168 subjects who completed the treatment phase of study 303 entered study 305 (60 LEF, 60 SSZ, 26 placebo/SSZ). 116 subjects completed 2 years of double-blind treatment in studies 301, 303, and 305 (49 LEF, 46 SSZ, 21 placebo/SSZ).
109 (74.7%) of the treated subjects were women and 37 (25.3%) were men (LEF: 82% women, SSZ: 68% women, placebo/SSZ: 73% women). 130 (89%) subjects were white. Mean (± SD) ages in the 3 treatment groups were similar (LEF: 57.8 ± 10.8 years, SSZ: 58.8 ± 11.4 years, placebo/SSZ: 58.6 ± 12.5 years). The treatment groups were comparable with regard to duration of RA and **age** at onset of RA.

Study regimen

(numbers in table show N (%) of subjects withdrawn before their 2-year cut-off visit)

Reason for withdrawal	LEF	**ssz**	**Plac./JSSZ**
Lack of efficacy	1 (1.7)	3 (5.0)	3(11.5)
Adverse events (incl. 2 deaths, LEF)	6(10.0)	9 (15.0)	1 (3.8)
Refusal/Noncompliance	3 (5.0)	1 (1.7)	0 (0)
Lost to follow-up	1 (1.7)	1 1.7)	1 (3.8)
Total withdrawals	**11 (18.3)**	**14 (23.3)**	**5 (19.2)**

Efficacy

65 of the 146 subjects were evaluable for the intention-to-treat analyses of total Larsen X-ray score of hands and forefeet (LEF 28, SSZ 27, placebo/SSZ 10). 25 LEF, 24 SSZ, and 8 placebo/SSZ subjects were included in the per-protocol analyses of total Larsen X-ray score.

The adjusted mean changes in total Larsen score, joint counts and global assessments over the 2-year treatment period between baseline (visit 0, study 301) and endpoint (study 305) in the intention-to-treat population were:

Variable	Mean (SD)				p-value
	LEF		SSZ		LEF vs
	Baseline	Adj. change	Baseline	Adj. change	ssz
Total Larsen X-ray score	1.43 (0.652)	-0.05 (0.215)	1.39 (0.592)	-0.04 (0.215)	0.6731
Tender joint count	18.4 (6.73)	-12.2 (6.69)	15.7 (5.48)	-10.5 (6.69)	0.1901
Swollen joint count	16.7 (5.75)	-9.9 (5.65)	15.2 (4.89)	-8.6 (5.65)	0.2266
Inv. global assessment	3.6 (0.64)	-1.5 (0.85)	3.4 (0.57)	-1.1 (0.85)	0.0344
Pat. global assessment	3.7 (0.69)	-1.6 (0.86)	3.5 (0.63)	-1 .o (0.86)	0.0007

Total Larsen X-ray score: LEF N=28, SSZ N=27; joint counts and global assessments: LEF N=60, SSZ N=57

The total Larsen X-ray score was used as the primary variable in study 303. Changes in this variable were very small, indicating virtually no deterioration in RA-related X-ray findings in these treatment groups. The results for the per-protocol population of subjects treated for at least 94 weeks were very similar.

Joint counts and global assessments were used as the primary efficacy variables in study 301. In both the LEF and SSZ groups, the effects achieved on these variables in the first 24 weeks of treatment (study 301) were well maintained on further treatment in the second 24-week treatment period (study 303) and in the second year of treatment (study 305). Similar findings were obtained for all other efficacy variables. Improvements in the LEF group between baseline (study 301) and endpoint (study 305) were significantly better than in the SSZ group for the investigators and patients global assessments. Responder rates were also significantly better in the LEF group (ACR: LEF 82%, SSZ 60%, p=0.0085; Paulus: LEF 87%, SSZ 72%, p=0.0430).

Safety

All 146 subjects were evaluated for safety. Treatment-emergent primary events were reported in study 305 for 88% of LEF subjects, 97% of SSZ subjects, and 89% of placebo/SSZ subjects. The most common treatment-emergent primary events observed on LEF (≥10% subjects) were RA (37%) upper respiratory infection (30%) rash (17%), arthralgia (15%) bronchitis (13%), cough increased (10%) joint disorder (10%) and dyspepsia (10%). The only possibly related treatment-emergent primary events to be reported in ≥10% of LEF subjects were RA (15%) and rash (10%). Rash, dyspepsia, and cough increased were more common on LEF than on SSZ.

No unexpected adverse events occurred in subjects who were treated with LEF for more than 1 year (i.e. in study 305). RA, upper respiratory infection, arthralgia, bronchitis, and joint disorder showed higher frequencies in study 305 than in study 301 in both LEF and SSZ subjects. These events are typical of the underlying disease and patient population, and higher frequencies can be expected due to the longer duration of treatment in study 305.

Frequencies of rash, dyspepsia, and cough increased were only higher in study 305 than in study 301 for the LEF group. In the SSZ group, frequencies of rash and cough increased were comparable in studies 305 and 301, but dyspepsia showed a lower frequency in study 305 than in study 301. The following events were less frequent in study 301 than in study 305: diarrhea, nausea, and headache in the LEF group; nausea, alopecia, and headache in the SSZ group.

Analysis over the entire treatment period (i.e. from start of study 301) showed that most first occurrences of diarrhea, hypertension, nausea, alopecia, abnormal liver function test, and pruritus in the LEF group took place in the first 3 months of treatment. In the SSZ group, most first occurrences of diarrhea, hypertension, nausea, and rash took place in the first 3 months of treatment. First occurrences of rash in the LEF group were more evenly distributed across the three time intervals studied (≤3 **months, 4-12 months, >12 months).**

52 (36%) of the 146 subjects had a total of 91 serious treatment-emergent primary events. The only serious events reported in more than 2 subjects of a treatment group were joint disorder (LEF 4 subjects, SSZ 4 subjects, placebo/SSZ 2 subjects) and RA (2 vs 2 vs 4 subjects). 4 subjects died (LEF 2, SSZ 1, placebo/SSZ 1). All deaths were probably of cardiogenic origin and considered unrelated to the study medication. 28 (19%) of the 146 subjects had treatment-emergent primary events that led to withdrawal of study medication before or after the 2-year cut off visit. The rate of withdrawal from LEF due to adverse events was lower (10 of 60 subjects) than for SSZ (15 of 60 subjects).

The increase in hemoglobin and decreases in white blood count and platelets on LEF reflect the therapeutic effect of the drug over the 2-year treatment period. Significant decreases in white blood count and platelets were also seen on SSZ but hemoglobin values did not change and hematocrit and erythrocyte values decreased. Significant increases in albumin and decreases in total protein also reflect effective disease control in both treatment groups. No clinically relevant changes were seen in any other laboratory variables.

No clinically relevant differences were detectable between the LEF and SSZ groups for pulse, body temperature, 1Zlead ECG, or chest X-ray. Mean blood pressure showed a small increase in the LEF group and a small decrease in the SSZ group. Body weight showed a small decrease in the LEF group and a small increase in the SSZ group.

COMMENTS/ CONCLUSIONS

In conclusion, hardly any deterioration in X-ray findings for hands and forefeet was observed in either the LEF or the SSZ group over the 2-year treatment period between baseline of study 301 and endpoint of study 305. In both treatment groups, the improvements achieved in all efficacy variables in the first 24 weeks (study 301) were well maintained during the rest of the 2-year treatment period. Overall, LEF was well tolerated during long-term treatment in subjects with RA. It cannot, however, be excluded that LEF may have a mild hypertensive effect and may lead to a decrease in body weight in some subjects.

Hoechst

STUDY SYNOPSIS HWA 486/6/MN/304/RA

TITLE	Comparative extension trial of the safety and efficacy of leflunomide versus methotrexate in patients with active rheumatoid arthritis
INVESTIGATORS Study Sites	Multinational / 97 centers: Belgium (2), Denmark (11), Finland (5) France (3), Germany (27), Hungary (1), Ireland (3), Netherlands (9), Norway (5), South Africa (2), Spain (3), Sweden (6), United Kingdom (20)
STUDY DATES	23 February 1995 - 20 February 1998
REPORT TYPE	Clinical/biometric report, final
REPORT ORIGIN	ClinData International (Pty) Ltd, Bloemfontein, South Africa Hoechst Marion Roussel Deutschland GmbH, Frankfurt, Germany
DATE OF ISSUE	5 November 1998
PHASE	III
INDICATION	Rheumatoid arthritis (RA)
STUDY OBJECTIVES	**Primary objective:** To investigate the safety of leflunomide (LEF) during long-term use in RA subjects. **Secondary objectives:** •To compare the efficacy and safety profiles of LEF with those of methotrexate (MTX) after long-term treatment. •To perform a pharmacoeconomic evaluation and investigate population *pharmacokinetics*.

STUDY MEDICATION AND DOSAGE			
	Drug	LEF	MTX
	Formulation	10 mg film-coated tablets	2.5 and 7.5 mg tablets
	Dosage	20 mg or 10 mg daily	7.5, 10 or 15 mg weekly
	Batch Nos.	See Informational Appendix B of study report	

	As the LEF and MTX treatments differed in appearance, placebos to both LEF and MTX were used in a double-dummy technique to preserve the double-blind nature of the trial.
STUDY DESIGN	Multinational, parallel-group study with a double-blind treatment phase and a 12-week, treatment-free observation phase. The study was an extension to the 52-week study HWA 486/6/MN/302/RA ("study 302"). The double-blind treatment period of study 304 continued until the last subject had completed 2 years of treatment in studies 302 and 304. Subjects who received LEF or MTX in study 302 continued on the respective medication in study 304. At the start of study 304, all subjects continued on the same dosage of LEF or MTX that they had been taking at completion of study 302.
STUDY POPULATION	Subjects with active RA who had completed 52 weeks of treatment in study 302 and who wished to continue treatment.

Study manager: Dr I Liiw-Friedrich
Authors: Dr I Lijw-Friedrich, Ch Oed, S Harris
Biometricians: Ch Oed, S Harris
Medical writer: Dr G Schulz

STUDY VARIABLES

Efficacy X-ray evaluation of hands and forefeet (total Larsen X-ray score was the main variable), tender joint count, swollen joint count, investigator's global assessment of RA activity, patient's global assessment of RA activity, responder rates (Paulus and American College of Rheumatology [ACR]), treatment success rate, joint tenderness score, swollen joint score, duration of morning stiffness, pain intensity assessment, Health Assessment Questionnaire (HAQ), erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), rheumatoid factor (RF)

Safety Observed and spontaneously reported adverse events; hematology, blood chemistry, urinalysis; physical examination, supine blood pressure, pulse, oral body temperature, body weight, 12lead electrocardiogram, chest X-ray

Pharmacokinetics / Pharmacoeconomics Trough blood samples for determination of plasma/serum levels of A77 726. Utilization of health-care resources.
Results for both analyses to be presented in separate reports.

STATISTICAL METHODS In addition to descriptive statistics, the following tests were performed:

Baseline comparability

Categorical variables: x2 test (if cell frequency below 5, Fisher's exact test) Continuous variables: analysis of variance (ANOVA)

Efficacy variables

X-ray evaluation of hands and forefeet (total Larsen score, subscores, number of eroded joints): analysis of covariance (ANCOVA). Responder and treatment success rates: logistic regression. All other variables: ANCOVA. The main efficacy analysis was the intention-to-treat comparison of the mean changes in total Larsen X-ray score for hands and forefeet from baseline in study 302 to endpoint in study 304 between LEF and MTX.

Safety variables

Time to withdrawal/time to first occurrence of events: life-table analysis Laboratory data: Wilcoxon signed rank test

RESULTS

Study sample 612 (83%) of the 736 subjects who completed the treatment phase of study 302 entered study 304 (292 LEF, 320 MTX). 497 subjects completed 2 years of double-blind treatment in studies 302 and 304 (233 LEF, 264 MTX).
436 (71%) of the treated subjects were women and 176 (29%) were men. 606 (99%) subjects were white. Mean ages (&SD) in the 2 treatment groups were similar (LEF: 57.7±9.8 years, MTX: 57.0±11.0 years). The treatment groups were comparable with regard to duration of RA and age at onset of RA.

Study regimen

No. (%) of subjects withdrawn before their 2-year cut-off visit

Reason for withdrawal	LEF	MTX
Did not want to continue	2(0.7)	5(1.6)
Adverse event	24(8.2j	19(5.9j
Unsatisfactory therapeutic response	17(5.8)	14(4.4)
Non-compliance	6(2.1)	6(1.9)
Lost to follow-up	4(1.4)	2(0.6)
Death	1(0.3)	7(2.2)
Protocol violation	1(0.3)	0(0.0)
Other reason	4(1.4)	3(0.9)
Total withdrawals	**59(20.2)**	**56(17.3)**

Efficacy 358 of the 612 subjects were evaluable for the intention-to-treat analyses of total Larsen X-ray score of hands and forefeet (LEF 167, MTX 191). 133 LEF and 158

MTX subjects were included in the per-protocol analyses of the total Larsen X-ray score.

The adjusted mean changes in total Larsen score, joint counts and global assessments over the 2-year treatment period between baseline (visit 0, study 302) and endpoint (study 304) in the intention-to-treat population were:

Variable	Mean(SD)				p-value LEF vs MTX
	LEF		MTX		
	Baseline	Adj. change	Baseline	Adj. change	
Total Larsen X-ray score	1.27 (0.474)	0.02 (0.223)	1.31 (0.521)	-0.03 (0.223)	0.0238
Tender joint count	16.9 (6.73)	-10.55 (6.76)	17.2 (6.50)	-10.89 (6.76)	0.5428
Swollen joint count	16.0 (6.03)	-9.10 (6.00)	16.1 (6.02)	-10.31 (6.00)	0.0173
Inv. global assessment	3.5 (0.58)	-1.17 (0.85)	3.6 (0.60)	-1.34 (0.85)	0.0147
1 Pat. global assessment	3.5 (0.65)	-1.24 (0.88)	3.6 (0.66)	-1.26 (0.88)	0.8072

Total Larsen X-ray score: LEF N=167, MTX N=191; joint counts and global assessments: LEF N=273, MTX N=298

The mean total Larsen X-ray score showed a slight increase (deterioration) after LEF treatment for 2 years. The MTX group showed a marginal decrease (improvement). The treament difference was statistically significant. The results for the per-protocol population of subjects treated for ~102 weeks were very similar.

Joint counts and global assessments were used as the primary efficacy variables in study 302. In both the LEF and MTX groups, the effects achieved on these variables in the first 52 weeks of treatment (study 302) were maintained on further treatment in the second 52-week treatment period (study 304). Similar findings were obtained for LEF and MTX for all other efficacy variables. Improvements in the MTX group between baseline and endpoint for the swollen joint count and the investigator's global assessment were significantly better than in the LEF group. Responder rates were higher in the MTX group (ACR: LEF 64%, MTX 72%, p= 0.0555; Paulus: LEF 71%, MTX 82%, p= 0.0007).

Safety

All 612 subjects were evaluated for safety. Treatment-emergent primary events were reported in study 304 for 90% of LEF subjects and 95% of MTX subjects, and possibly related primary events in 55% LEF subjects and 50% MTX subjects. The most common treatment-emergent primary events observed on LEF (210% of subjects) in study 304 were upper respiratory infection (32%), RA (26%) joint disorder (14%) bronchitis (13%) back pain (13%) hypertension (12%) rash (11%) accidental injury (IO%), and diarrhea (10%). The most common primary events on MTX were upper respiratory infection (38%), RA (26%), bronchitis (13%), and accidental injury (12%).

The most common possibly related treatment-emergent primary events in LEF subjects were hypertension (7%) rash (7%), upper respiratory infection (5%) and alopecia (5%). The most common possibly related events in the MTX group were upper respiratory infection (6%) liver function test abnormal (6%) and alopecia (5%). No unexpected adverse events occurred in subjects who were treated with LEF for more than 1 year (i.e. in study 304). Upper respiratory infection, RA, joint disorder, bronchitis, back pain, and accidental injury showed higher frequencies in study 304 than in study 302 in both LEF and MTX subjects. These events are typical of the underlying disease and patient population, and higher frequencies can be expected due to the longer duration of treatment in study 304 (1 year in 302 vs up to 3 years in 304). Frequencies of some of the adverse events that are known to be associated with the respective drugs showed lower frequencies in the second year of treatment (study 304) than in the first year of treatment (study 302). Alopecia, nausea, and abnormal liver function test were less common in study 304 in both treatment groups. Diarrhea was less common in study 304 in the LEF group but not in the MTX group Analysis over the entire treatment period (i.e. from start of study 302) showed that, in both treatment groups, most first

occurrences of diarrhea and nausea took place in the first 3 months of treatment and most first occurrences of liver function test abnormal and alopecia took place in the first year of treatment. First occurrences of hypertension, abdominal pain, and rash were more evenly distributed across the three time intervals studied (2 3 months, 4-12 months, >12 months).

214 (35%) subjects had a total of 365 serious treatment-emergent primary events.

The most common types of serious events were joint disorder (27 LEF vs. 19 MTX subjects), RA (19 vs 23) and synovitis (15 vs 5). 13 subjects (2 LEF, 11 MTX) died. Deaths were considered possibly related to the study medication in 2 MTX subjects (pancytopenia, pneumonitis). In the LEF group, 1 death was associated with myocardial infarction, and 1 with carcinoma. In the MTX group, 2 deaths were associated with drug-related pneumonitis, 2 with pneumonia, 5 with cerebra-/ cardiovascular disorders, 1 with carcinoma, and 1 with septic shock. 62 (10%) of the 612 subjects (32 LEF, 30 MTX) had treatment-emergent primary events that led to withdrawal of study medication before or after the 2-year cut off visit.

Increases in hemoglobin and erythrocyte count and decreases in white blood count and platelets on LEF reflect the therapeutic effect of the drug over the 2-year treatment period. Comparable results were obtained in the MTX group. Significant increases in albumin and decreases in total protein also reflect effective RA control in both treatment groups. SGOT and SGPT increased in both treatment groups. Frequencies of clinically noteworthy changes in SGPT were higher on MTX than LEF (13 vs 6 subjects), as were abnormal liver function tests reported as primary adverse events in study 304 (22 vs 9). No clinically relevant changes were seen in any other laboratory variables.

Over the 2-year treatment period, no clinically relevant differences were detectable between the LEF and MTX groups for pulse, body temperature, 12-lead ECG, or chest X-ray. Mean blood pressure showed a small increase in both treatment groups with a larger increase in the LEF group; hypertension was more commonly reported as an adverse event on LEF than on MTX (36 vs. 18 subjects). Body weight showed a small decrease in the LEF group and a small increase in the MTX group. The treatment differences for body weight and hypertension at endpoint were statistically significant.

COMMENTS CONCLUSIONS

In conclusion, only small changes in X-ray findings for hands and forefeet were observed in the leflunomide and methotrexate groups, indicating effective disease control. The improvements achieved in all efficacy variables in the first year of treatment (study 302) were well maintained in the second year of treatment (study 304) in both the leflunomide and methotrexate groups. Although improvements under leflunomide were slightly lower than those under methotrexate, the differences between the two treatment groups are not clinically meaningful. The overall improvement in both groups is considered clinically relevant. Overall, leflunomide was well tolerated during long-term treatment in subjects with RA.

STUDY SYNOPSIS
HWA 486/F/USA/301/RA
ULTRA Study (Utilization of Leflunomide for Treatment of Rheumatoid Arthritis)
2 Year and Alternate Therapy

TITLE

A Phase III, Double Blind, Randomized, Placebo-Controlled Study to Compare the Activity and Safety of Leflunomide to Methotrexate or Placebo in Subjects with Active Rheumatoid Arthritis (Final 24-Month Data Including Alternate Therapy)

INVESTIGATORS, STUDY SITES

Multinational/47 centers: United States of America (42), Canada (5)

STUDY DATES

May 30, 1995 – November, 1998

REPORT TYPE

Clinical/biometric report, final.
This is the second final study report for this protocol. The first study report included results from the first year of therapy, as this was the primary efficacy endpoint and was used to support applications submitted globally for approval of leflunomide. The current report presents data from two years of therapy (the total time frame of the protocol) and from the alternate therapy phase of the protocol).

REPORT ORIGIN

Quintiles, San Diego, CA, USA
Hoechst Marion Roussel Deutschland GmbH, Frankfurt, Germany

DATE OF ISSUE

August 25, 1999

PHASE

III

INDICATION

Rheumatoid arthritis (RA)

STUDY OBJECTIVES

Primary objectives. To compare the efficacy and safety of leflunomide (LEF) with placebo (PBO) in subjects with active RA who had never previously received methotrexate (MTX).

Secondary objectives. To compare the efficacy and safety of LEF with MTX in subjects with active RA.

STUDY DESIGN

Phase III, multinational, multicenter, double-blind, randomized, placebo- and MTX-controlled study of parallel group design

STUDY MEDICATION AND DOSAGE

	LEF (once per day)	PBO	MTX (once per week)
Drug formulation/ Frequency	100 mg, Days 1-3 20 mg, Days 4-7 20 mg, Wks 2-6 20 mg if tolerated*, Wks 7-104	Matching placebo dispensed at times scheduled for LEF or MTX	7.5 mg, Wks 1-6 7.5 or 10 mg**, Wk 7 7.5 or 12.5 mg**, Wk 8 7.5 or 15 mg**, Wks 9-52 7.5, 15, 17.5 or 20 mg, Wks 53-104

* Dose could be lowered to 10 mg/day depending on tolerance.
** If active disease was present at Week 6.
All subjects were to receive folate (1 mg, twice daily).

Batch Nos.

See Appendix **B** of study report

DURATION OF

TREATMENT Up to 104 weeks

STUDY POPULATION Subjects with a diagnosis of RA by ACR criteria of ≥ 6 months' duration, who had active RA by ACR criteria at screening and baseline

STUDY VARIABLES

Efficacy

Primary. ACR20 responder rate: Percentage of subjects who met the ACR20 responder criteria at endpoint.

Secondary. Tender joint count, swollen joint count, patient global assessment, physician global assessment, modified Health Assessment Questionnaire (MHAQ), pain intensity assessment, erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), morning stiffness, X-rays of hands and feet, questionnaires on quality of life and functional ability (Medical Outcomes Study (MOS) SF-36, MOS current health perceptions scale, Work Limitations Questionnaire, standard HAQ, and Problem Elicitation Technique (PET))

Safety Adverse events; hematology, blood chemistry, urinalysis; physical examination, supine blood pressure, heart rate, oral body temperature, body weight, 12-lead ECG, and chest X-ray.

Pharmacokinetics Blood samples for determination of plasma concentrations of A77 1726 and TFMA. Results will be presented in a separate report.

Pharmacoeconomics Utilization of health-care resources; studied for first 52 weeks only. Results were presented in a previous report.

STATISTICAL METHODS The statistical analyses were performed for three cohorts of subjects: (a) intent-to-treat cohort (ITT cohort), (b) the subset of subjects with data during year 2 of therapy (year-2 cohort), and (c) subjects enrolled in the alternate therapy phase. In addition to descriptive statistics, the following tests were performed:

Efficacy variables Primary: ACR20 responder rates at endpoint in the ITT cohort analyzed by logistic regression.
Secondary: All secondary efficacy variables: analysis of covariance (ANCOVA) for all those cohorts analyzed.

RESULTS

Study sample

In total, 511 subjects were randomized; 3 were randomized but not treated, 508 were randomized and treated (190 LEF, 128 PBO, and 190 MTX); 235 completed greater than 52 weeks of therapy (98 LEF, 36 PBO, and 101 MTX); 116 entered the alternate therapy 1-year treatment group (35 MTX to LEF, 56 PBO to LEF, and 25 LEF to MTX).

Of the treated subjects in the ITT cohort, 371 (73.0%) were females and 137 (27.0%) were males: LEF 72.6% females; PBO 71.9% females; MTX 74.2% females. Mean ages (years) in the 3 treatment groups were similar: LEF 54.0±12.1, PBO 54.7±10.6, and MTX 53.3±11.6. The treatment groups were similar for duration of RA, age at onset of RA, and prior disease modifying antirheumatic drug (DMARD) use.

Of the treated subjects in the year-2 cohort, 162 (68.9%) were females and 73 (31.1%) were males: LEF 69.4% females, PBO 69.4% females, and MTX 68.3% females. Mean ages (years) in the 3 treatment groups were similar: LEF 55.2±11.7, PBO 54.2±11.4, and MTX 53.3±12.4. The treatment groups were similar for duration of RA, age at onset of RA, and prior DMARD use.

Of the treated subjects in the alternate therapy group, 90 (77.6%) were females and 26 (22.4%) were males: PBO/LEF 75.0% females, MTX/LEF 88.6% females, LEF/MTX 68.0% females. Mean ages (years) in the 3 treatment groups were similar: PBO/ LEF 56.7±11.1, MTX/LEF 52.4±10.8, and LEF/MTX 52.2±12.5.

Study regimen

Reasons for early withdrawal from the ITT cohort

Reason	LEF (N=190)		PBO (N=128)		MTX (N=190)		Total (N=508)
	N	%	N	%	N	%	
Lack of efficacy	37	**19.5**	71	**55.5**	54	**28.4**	162
Adverse event	52	**27.4**	12	**9.4**	30	**15.8**	94
Lost to follow-up	1	**0.5**	1	**0.8**	4	**2.1**	6
Protocol violation	2	**1.1**	2	**1.6**	1	**0.5**	5
Noncompliance	–	**–**	1	**0.8**	2	**1.1**	3
Death	–	**–**	–	**–**	2	**1.1**	2
Other	15	**7.9**	14	**10.9**	17	**8.9**	46
Total	107	**56.3**	101	**78.9**	110	**57.9**	318

Reasons for early withdrawal from the year-2 cohort

Reason	LEF (N=98)		PBO (N=36)		MTX (N=101)		Total (N=235)
	N	%	N	%	N	%	
Lack of efficacy	4	**4.1**	2	**5.6**	5	**5.0**	11
Adverse event	8	**8.2**	–	**–**	8	**7.9**	16
Lost to follow-up	–	**–**	1	**2.8**	2	**2.0**	3
Protocol violation	1	**1.0**	1	**2.8**	–	**–**	2
Noncompliance	–	**–**	–	**–**	1	**1.0**	1
Death	–	**–**	–	**–**	1	**1.0**	1
Other	2	**2.0**	5	**13.9**	4	**4.0**	11
Total	15	**15.3**	9	**25.0**	21	**20.8**	45

Efficacy

506 subjects were included in the ITT cohort (LEF 190, PBO 128, MTX 188); 235 subjects were included in the year-2 cohort (LEF 98, placebo 36, MTX 101).

LEF vs placebo. In the ITT cohort, ACR20 response rates at Month 24 (by LOCF) showed leflunomide to be statistically significantly better than placebo. Analyses of mean changes in the components of the ACR20 response over time showed leflunomide to be highly statistically superior to placebo for all measures.

Month 24 ACR20 response rates (LOCF): ITT Cohort

Treatment	No. of Subjects ACR20 Responders/Total	ACR20 Response Rate	p-value (95% confidence interval)
LEF	99/186	53.2	LEF vs PL: p≤0.001, 95% CI 21.7% to 41.9
PBO	34/128	26.6	LEF vs MTX: p=0.317, 95% CI-5.0% to 15.3%
MTX	90/188	47.9	MTX vs PL: p≤0.001, 95% CI 16.7% to 36.7%

In the year-2 cohort, ACR20 response rates were high in all groups since the year-2 cohort was enriched by subjects with positive treatment effect continuing into a second year of therapy. The size of the placebo group (36, 28% of the originally enrolled subjects) prohibited statistical comparisons with that group. Analyses of mean changes in the components of the ACR20 response over time showed similarly high improvement in the signs and symptoms of RA in all treatment groups, with

the greatest relief seen in the leflunomide group. The year-2 cohort results indicate sustained effect over two years of treatment.

12 and 24 Month ACR20 response rates (LOCF): Year-2 Cohort

Treatment	12 Month No. of ACR20 Responders/Total	ACR20 Response Rate	24 Month No. of ACR20 Responders/Total	ACR20 Response Rate	p-value (95% confidence interval)
LEF	75/97	77.3	77/97	79.4	LEF vs MTX:
PBO	22/36	61.1	23/36	63.9	p=0.049, 95% CI
MTX	61/101	60.4	68/101	67.3	0.1% to 24.4%

Statistical methods: Logistic regression: 95% CI for LEF vs MTX.

Analyses of the Sharp scores for X-rays of hands and feet for the ITT cohort at 12 months showed that increase in the total Sharp score was statistically significantly lower in the leflunomide group than the placebo group (0.5 vs 1.9, p=0.0016) This was also demonstrated in joint-space-narrowing subscore; no difference for the mean change in the erosion subscore was found. For the year-2 cohort, the Sharp scores showed that in subjects who continued into year 2, little progression was seen in all treatment groups at both year 1 and year 2, indicating continued protection against joint deterioration.

Analyses of the functional ability and health-related quality of life measures showed that the use of LEF for up to 104 weeks maintained improvements in physical function and health-related quality of life demonstrated after one year of treatment.

LEF vs MTX. In the ITT cohort, the ACR20 response rates for the LEF and MTX groups were statistically equivalent. The ACR20 response rate for LEF was 53.2% and for MTX was 47.9%. Response rates over time showed the LEF response occurred earlier and was higher than in either of the other two treatment groups. The MTX response was statistically significantly better than the placebo response. In the year-2 cohort, the ACR20 response rates for the LEF and MTX groups were 79.4% and 67.3% respectively, which were not statistically equivalent.

The X-ray analysis of hands and feet for the ITT cohort showed LEF and MTX to be statistically equivalent. The mean increase in total Sharp score was statistically significantly lower in the MTX group than in the placebo group. For the year-2 cohort, the results showed similar and low increases in total Sharp scores in both groups over year 2.

Functional ability and health-related quality of life measures showed better responses in the LEF group than in the MTX group on multivariate analysis, demonstrating overall statistical significance.

Safety

Adverse events. In the ITT cohort, there were 84 (16.5%) serious adverse events; they were reported more frequently in the LEF (18.9%) and MTX (18.9%) than in the PBO (9.4%) groups. Few of these events were considered related to study drug (LEF 1.6%, PBO 1.6, and MTX 3.7%). In the year-2 cohort, there were 53 (22.6%) serious adverse events; they were reported more frequently in the MTX (24.8%) than the LEF (22.4%) and PBO (16.7%) groups. Few of these events were considered related to study drug (LEF 1.0%, MTX 2.0%). There were no deaths on study in the LEF, 2 in the MTX (1 study drug related), and 1 in the PBO groups.

Withdrawals due to serious adverse events were lower in the LEF compared to the MTX group in both the ITT and year-2 cohorts (LEF 4.2%, PBO 1.6%, and MTX 6.3%), and (LEF 1.0%, PBO 0%, and MTX 5.0%), respectively. Serious adverse events which occurred in > 1% LEF subjects in the ITT cohort were pneumonia (4 subjects), infection (2)**,** joint disorder (5), cholelithiasis (4), hypertension (2), and deep thrombophlebitis (2); for the year-2 cohort, they were cholelithiasis (4), joint disorder (4), infection (2), hypertension (2), cholecystitis (2), and deep thrombophlebitis (2).

Adverse events (i.e., serious and non-serious events) in the ITT cohort were more frequent in the LEF (98.4%) than in the PBO (88.3%) or MTX (92.1%) groups. In the year-2 cohort, the frequency of adverse events was similar for all treatment groups: LEF (100.0%), PBO (94.4%), and MTX (96.0%). The most frequently reported adverse events, regardless of causality, in the LEF group in the ITT cohort, were: respiratory infection (37.4%), diarrhea (36.8%), headache (20.0%), hypertension (18.4%), and rash (17.4%). In the PBO group: respiratory infection (25.0%), nausea (18.8%), headache (17.2%), and dyspepsia (16.4%). In the MTX group: respiratory infection (38.4%), headache (23.2%), diarrhea (21.6%), and nausea (20.5%). In the year-2 cohort, the most frequently reported adverse events, regardless of causality, in the LEF group were: respiratory infection (55.1%), diarrhea (43.9%), hypertension (28.6%), dyspepsia (24.5%), and rash (21.4%). In the PBO group: respiratory infection (38.9%), diarrhea (30.6%), dyspepsia (27.8%), and accidental injury (25.0%). In the MTX group: respiratory infection (54.5%), diarrhea (23.8%), headache (23.8%), and accidental injury (20.8%).

In the ITT cohort, adverse events considered related to study drug occurred in LEF (80.5%), PBO (56.3%), and MTX (65.8%) of subjects. The frequency of withdrawals due to adverse events was higher in the LEF (26.8%) than in the PBO (9.4%) or MTX (16.8%) groups. For the year-2 cohort, adverse events considered related to study drug occurred more frequently in the LEF (81.6%), compared to the PBO (63.9%), and MTX (62.4%) groups. The frequency of withdrawals due to adverse events was similar in the LEF (7.1%) compared to the MTX (8.9%) group, with PBO (0%).

In the ITT cohort, the most common adverse events considered related to LEF administration were of gastrointestinal origin, predominantly diarrhea (27.9% LEF, 13.3% PBO, and 13.7% MTX), LFT abnormalities (15.3% LEF, and 10.5% MTX), dyspepsia (LEF 13.2%, PBO 13.3%, and MTX 8.9%) and nausea (LEF 12.6%, PBO 18.0%, and MTX 15.8%). In the year-2 cohort, the most common adverse events considered related were diarrhea (LEF 31.6%, PBO 22.2%, and MTX 11.9%), dyspepsia (LEF 18.4%, PBO 16.7%, and MTX 8.9%), rash (LEF 13.3%, PBO 2.8%, and MTX 3.0%), alopecia (LEF 13.3%, PBO 0%, and MTX 4.0%), hypertension (LEF 12.2%, PBO 5.6%, and MTX 1.0%), abdominal pain, digestive system (LEF 10.2%, PBO 5.6%, and MTX 7.9%), and LFT abnormalities (LEF 10.2%, PBO 2.8%, and MTX 8.9%).

Of the 63 total subjects in the ITT cohort with adverse events leading to withdrawal from the study which were related, 38 (20.0%) received LEF, 7 (5.5%) PBO, and 18 (9.5%) MTX. Of these, the most frequent events in the LEF group (21 subjects, 11.1%) were in the digestive system, and included LFT abnormalities, diarrhea, and nausea. Of the 6 total subjects in the year-2 cohort with adverse events leading to withdrawal from the study which were related, 4 (4.1%) received LEF, 0% PBO, and 2 (2.0%) MTX.

A comparison of the incidence of adverse events in year 1 with year 2 in the year-2 cohort showed that, in general, adverse events decreased in year 2 in all treatment groups. Exceptions were hypertension increased in LEF from 15.3% to 23.5%, PBO from 5.6% to 11.1%, and MTX from 3.0% to 4.0%; arthralgia increased in LEF from 5.1% to 11.2%; and peripheral edema increased in LEF from 5.1% to 10.2%.

The highest incidence of grouped adverse events in the leflunomide treatment group in the ITT cohort was reported in the gastrointestinal system. Diarrhea was the most frequently reported: LEF (36.8%), PBO (20.3%), and MTX (21.6%). The majority of diarrhea adverse events in the LEF group were mild to moderate, all resolved without sequelae, and occurred during the first 1-2 months of study drug administration, an effect that may have been related to the loading dose of leflunomide at the initiation of study. Study treatment was decreased in 5 (2.6%) LEF subjects; interrupted in 3 (1.6%) LEF, 1 (0.8%) PBO, and 6 (3.2%) MTX subjects; and discontinued in 18 (9.5%) LEF, 3 (2.3%) PBO, and 10 (5.3%) MTX subjects. In the year-2 cohort, diarrhea was also the most frequently reported: LEF (43.9%), PBO (30.6%), and MTX (23.8%). These adverse events in the LEF group were mild to moderate, and, with the exception of 2 (2%) subjects, all resolved. Study treatment was decreased in 4 (4.1%) LEF, and 1 (1.0%) MTX subjects; interrupted in 2 (2.0%) LEF and 2 (2.0%) MTX subjects; and discontinued in 6 (6.1%) LEF and 3 (3.0%) MTX subjects. Diarrhea appeared to be an adverse effect of LEF, however, in both cohorts the majority of the diarrhea was mild to moderate.

In the ITT cohort, nausea/vomiting was reported at a similar rate for the three treatment groups: LEF (23.7%), PBO (22.7%), and MTX (21.6%). The majority of cases were mild to moderate, and 5 (2.6%) LEF and 5 (2.6%) MTX subjects had severe cases. Study treatment was decreased in 1 (0.5%) LEF and 1 (0.5%) MTX subjects; interrupted in 4 (2.1%) LEF, 1 (0.8%) PBO, and 6 (3.2%) MTX subjects; and discontinued in 19 (10.0%) LEF, 2 (1.6%) PBO, and 10 (5.3%) MTX subjects. In the year-2 cohort, nausea/vomiting was reported at a slightly lower rate in the LEF (18.4%) compared to PBO (25.0%) and MTX (20.8%) groups. Most cases were mild to moderate in severity and 2 (2.0%) LEF, and 1 (2.8%) PBO subjects required treatment. Study treatment was decreased in 1 (1.0%) LEF and 1 (1.0%) MTX subjects; interrupted in 1 (1.0%) MTX subject; and discontinued in 3 (3.1%) LEF and 4 (4.0%) MTX subjects. The incidence of nausea/vomiting and dyspepsia in both the ITT and year-2 cohorts were similar among treatment groups and may have been associated with NSAID administration. Nausea/vomiting did not appear to be an adverse effect of leflunomide administration.

Abdominal pain in the ITT cohort occurred at a slightly higher rate in LEF (18.9%) compared to PBO (10.9%) and MTX (14.7%) groups, and in the year-2 cohort, LEF (23.5%), PBO (22.2%), and MTX (16.8%). The similar occurrence of abdominal pain in all treatment groups in both cohorts may reflect NSAID use. Abdominal pain did not appear to be an adverse effect of leflunomide administration in both cohorts.

Oral ulcerations are an expected adverse event with methotrexate administration and occurred less frequently in the LEF group compared to the MTX group in the ITT cohort: LEF (6.8%), PBO (5.5%), and MTX (10.5%). In the year-2 cohort, the frequency was LEF (9.2%), PBO (2.8%), and MTX (14.9%).

In the ITT cohort, infections accounted for the second highest incidence of adverse events in all three treatment groups. Their occurrence was similar in the LEF (64.2%) and MTX (65.8%) groups, compared to the PBO (51.6%) group. The infections in the LEF group were generally mild to moderate. Treatment-related infections occurred in LEF (3.2%), MTX (2.1%), and PBO (0.8%) subjects. Study treatment was decreased in 5 (2.6%) LEF, and 6 (3.2%) MTX; interrupted in 6 (3.2%) LEF and 3 (1.6%) MTX subjects; and discontinued in 24 (12.6%) LEF, 5 (3.9%) PBO, and 18 (9.5%) MTX subjects. The most common infections were respiratory and their occurrence was similar in the LEF and MTX groups (37.4% and 38.4%, respectively), compared to PBO (25.0%) group. Respiratory infections related to study treatment were LEF (3.2%), MTX (2.1%), and PBO (0.8%). In the year-2 cohort, infections accounted for the highest incidence of grouped adverse events in all three treatment groups. Their occurrence was similar in the LEF (84.7%) and MTX (86.1%) groups, compared to the PBO (69.4%) group. Study treatment was decreased in 4 (4.1%) LEF and 6 (5.9%) MTX subjects; interrupted in 5 (5.1%) LEF and 3 (3.0%) MTX subjects; and discontinued in 7 (7.1%) LEF and 8 (7.9%) MTX subjects. The most common infections were respiratory and their occurrence was similar in the LEF and MTX groups (55.1% and 54.5%, respectively), compared to PBO (38.9%). Respiratory infections related to study treatment were LEF (5.1%), MTX (4.0%), and PBO (0%). There were no opportunistic infections or disseminated herpes zoster or herpes simplex.

Adverse events associated with the cardiovascular system in the ITT cohort were more frequently reported in the LEF (27.9%), compared to PBO (15.6%) and MTX (11.1%) groups. The most frequently reported adverse events were hypertension (LEF 18.4%, PBO 8.6%, and MTX 4.7%), and chest pain (LEF 7.9%, PBO 6.3%, and MTX 4.7%). The majority of these events in the LEF group were mild to moderate. Hypertension adverse events were related to study drug in 8.9% LEF, 4.7% PBO, and 0.5% MTX subjects. There were 2 (1.1%) serious adverse events associated with hypertension in the LEF, 1 (0.8%) PBO, and none in the MTX groups. Hypertension at baseline was reported at a higher frequency in the LEF subjects (13.7%), compared to PBO (8.6%), and MTX (2.1%) subjects. New-onset hypertension was reported in 4.7% LEF, 0% PBO, and 2.6% MTX subjects. In the year-2 cohort, cardiovascular adverse events were more frequently reported in the LEF (38.8%), compared to PBO (25.0%) and MTX (9.9%) groups. The most frequently reported adverse events were hypertension (LEF 28.6%, PBO 13.9%, and MTX 5.9%), and chest pain (LEF 9.2%, PBO 11.1%, and MTX 4.0%). The majority of these events in the LEF group were mild to moderate. Hypertension adverse events were possibly or probably related to study drug in 12.2% LEF, 5.6% PBO, and 1.0% MTX subjects. There were 2 (2.0%) serious adverse events associated with hypertension in the LEF group. Hypertension at baseline was reported in 21.4% LEF, 13.9% PBO, but only 2.0% MTX subjects. New-onset hypertension was reported in 7.1% LEF, 0% PBO, and 4.0% MTX subjects. In both the ITT and year-2 cohorts, hypertension at baseline and concomitant NSAID and steroid use were higher in the LEF compared to the PBO and MTX groups, which may reflect the higher occurrence of hypertension in the LEF group. Leflunomide administration did not appear to have any clinically significant effect on blood pressure in the ITT and year-2 cohorts.

In the ITT cohort, potential allergic reactions were reported in 29.5% LEF, 16.4% PBO, and 22.1% MTX subjects. The most commonly reported adverse events were rash (17.4% LEF, 8.6% PBO, and 11.1% MTX) and

allergic reactions (10.5% LEF, 4.7% PBO, and 6.3% MTX). Potential allergic reactions related to leflunomide administration consisted mainly of pruritus (3.7%) and rash (1.6%). Study treatment was decreased in 3 (1.6%) LEF and 2 (1.1%) MTX subjects; interrupted in 2 (1.1%) LEF, 1 (0.8%) PBO, and 2 (1.1%) MTX subjects**.** There were no serious adverse events. Treatment was discontinued for rash in 3 (1.6%) LEF, 3 (2.3%) PBO, and 0% MTX subjects. In the year-2 cohort, potential allergic reactions were reported in 38.8% LEF, 13.9% PBO, and 29.7% MTX subjects. The most commonly reported adverse events were rash (21.4% LEF, 8.3% PBO, and 12.9% MTX) and allergic reactions (17.3% LEF, 8.3% PBO, and 8.9% MTX). Study treatment was decreased in 3 (3.1%) LEF and 2 (2.0%) MTX subjects; interrupted in 1 (1.0%) LEF subject; and discontinued in 4 (4.1%) LEF and 1 (1.0%) MTX subjects. Potential allergic reactions related to leflunomide administration were higher than in the ITT cohort and consisted mainly of rash (13.3%), and pruritus (5.1%). No anaphylactic reactions or angioedema were noted in the ITT and year-2 cohorts.

In the ITT cohort, the incidence of rheumatoid arthritis (RA)-related adverse events were evenly distributed among the three treatment groups (26.3% LEF, 23.4% PBO, and 28.4% MTX). The majority of events were mild to moderate and were unrelated to study treatment. Vasculitis occurred in 1 (0.5%) LEF subject in year 1, which was unrelated to study drug administration, and 1 (0.5%) MTX subject in year 1, which was judged related to study drug administration. In the year-2 cohort, the incidence of RA-related adverse events was slightly higher than in the ITT cohort in the two active treatment groups (34.7% LEF and 38.6% MTX, compared to 25.0% PBO).

In the ITT cohort, the incidence of central nervous system adverse events was similar in all treatment groups (35.3% LEF, 28.1% PBO, and 35.3% MTX). Study treatment was decreased in 3 (1.6%) LEF, and 1 (0.5%) MTX subjects; interrupted in 1 (0.5%) LEF and 1 (0.5%) MTX subjects; and discontinued in 17 (8.9%) LEF and 17 (8.9%) MTX subjects. The most frequently reported event was headache (20.0% LEF, 17.2% PBO, and 23.2% MTX), of which related events occurred in 12.1% LEF, 7.8% PBO, and 12.6% MTX. The occurrence of grouped adverse events of neuritis/neuropathy/paresthesia was similar in the two active treatment groups (7.9% LEF and 7.4% MTX). In the year-2 cohort, the incidence of central nervous system adverse events was slightly higher than in the ITT cohort, and was similar in all treatment groups (43.9% LEF, 38.9% PBO, and 45.5% MTX). Study treatment was decreased in 2 (2.0%) LEF and 1 (1.0%) MTX subjects; and discontinued in 5 (5.1%) LEF and 5 (5.0%) MTX subjects. The most frequently reported event was headache (19.4% LEF, 19.4% PBO, and 23.8% MTX), of which related events occurred in 10.2% LEF, 8.3% PBO, and 10.9% MTX. The occurrence of grouped adverse events of neuritis/neuropathy/paresthesia was similar in the two active treatment groups 13.3% LEF and 8.9% MTX).

Other adverse events of interest included alopecia, which occurred in the ITT cohort in 10.5% LEF, 0.8% PBO, and 5.8% MTX subjects. In the year-2 cohort, alopecia occurred in 13.3% LEF and 5.0% MTX subjects. Most cases were mild to moderate and resolved without treatment. There were 9 discontinuations from study treatment due to alopecia in the ITT cohort (3 LEF, 1 PBO, and 5 MTX subjects). In the year-2 cohort, there was only 1 discontinuation due to alopecia in the LEF and 1 decrease in dosage in the MTX group. It should be noted that the alopecia resolved in the LEF subjects who discontinued. Alopecia occurred with leflunomide and

methotrexate administration, however the incidence in the LEF group declined from 12.2% in year 1 to 5.1% in year 2, compared to 5.0% and 3.0%, respectively for the MTX group.

Laboratory variables. Leflunomide administration in both the ITT and year-2 cohorts did not appear to be associated with clinically significant changes in hemoglobin, hematocrit, RBC parameters, platelets, and WBC subpopulations. There was no clinically significant effect on sodium, potassium, chloride, bicarbonate, BUN or creatinine. Leflunomide administration did appear to be associated with a decrease in serum uric acid in both the ITT and year-2 cohorts, due to the known uricosuric effect of leflunomide on the brush border membrane of the proximal renal tubule cells. In both the ITT and year-2 cohorts, leflunomide did not appear to be associated with clinically significant changes in total protein, albumin, and total bilirubin.

Leflunomide administration appeared to be associated with elevations of SGPT (ALT) and SGOT (AST) in the ITT and year-2 cohorts. In the ITT cohort, there were 10 (5.3%) LEF subjects with SGOT (AST) > 2x ULN to 3x ULN, and all normalized to ≤ 1.2x ULN. Similarly, of 12 (6.3%) LEF subjects with SGPT (ALT) > 2x to 3x ULN, all reversed to ≤ 2x ULN, 11 (5.8%) normalized to ≤ 1.2x ULN, 3 (1.6%) normalized after discontinuation, and 8 (4.2%) normalized without dose reduction of study treatment. There were 7 (3.7%) LEF subjects with SGOT (AST) > 3x ULN, and all normalized to ≤ 1.2x ULN. Similarly, of 12 (6.3%) LEF subjects with SGPT (ALT) > 3x ULN, all reversed to ≤ 2x ULN, 11 (5.8%) normalized to ≤ 1.2x ULN, 6 (3.2%) normalized after discontinuation, and 5 (2.6%) normalized without dose reduction of study treatment. Liver function abnormalities that were judged related to study drug were similar in the two active treatment groups, and occurred in 30 (15.8%) LEF, 4 (3.1%) PBO, and 22 (11.6%) MTX subjects. Of these, 14 (7.4%) LEF, 1 (0.8%) PBO, and 7 (3.7%) MTX subjects discontinued from the study. In the year-2 cohort, SGPT (ALT) was moderately elevated at any visit in 8.2% LEF, 5.6% PBO, and 5.0% MTX subjects; and marked elevations occurred in 6.1% LEF, 5.6% PBO, and 4.0% MTX subjects. SGOT (AST) was moderately elevated in 6.1% LEF, 0% PBO, and 5.0% MTX subjects; marked elevations occurred in 3.1% LEF, 2.8% PBO, and 1.0% MTX subjects; and elevations > 8x ULN occurred in 1.0% LEF subjects. At worst evaluation, there were 5.1% LEF subjects with SGOT (AST) > 2x ULN to 3x ULN, and all normalized to ≤ 1.2x ULN. Similarly, of 5.1% LEF subjects with SGPT (ALT) > 2x ULN to 3x ULN, all normalized without dose reduction. There were 4.1% LEF subjects with SGOT (AST) > 3x ULN, and all normalized to ≤ 1.2 x ULN. Similarly, of 6 (6.1%) LEF subjects with SGPT (ALT) > 3x ULN, all normalized to ≤ 1.2x ULN without dose reduction. Liver function abnormalities that were judged related to study drug were very similar in the two active treatment groups, and occurred in 10 (10.2%) LEF, 1 (2.8%) PBO, and 9 (8.9%) MTX subjects. Of these, 1 (1.0%) LEF subject discontinued from the study.

Leflunomide administration did not appear to be associated with any clinically significant adverse effect on alkaline phosphatase in the ITT and year-2 cohorts.

In both the ITT and year-2 cohorts, there were no adverse effects of leflunomide administration on other chemistry parameters: LDH, triglycerides, total cholesterol, calcium, phosphorous, glucose, and creatine kinase. In both the ITT and year-2 cohorts, there was essentially no difference between the three treatment groups in the tested parameters of urinalysis.

Clinical variables. In both the ITT and year-2 cohorts, there were no clinically significant differences between the three treatment groups in ECG, chest X-ray or physical examination results. Leflunomide administration had no effect on body temperature, weight, or heart rate.

Alternate Therapy Cohort

Efficacy

114 subjects were evaluable for efficacy in the alternate therapy phase of the protocol (56 PBO/LEF, 34 MTX/LEF, 24 LEF/MTX). Results were variable over time, but indicated that improvement occurred in all three treatment groups, not just those subjects entering from the placebo arm of initial therapy. Up to half of the subjects not responding to LEF or MTX responded well to the other DMARD. The percentage of subjects who were ACR20 responders at endpoint was somewhat higher in the PBO/LEF (52.8%) group than in the groups switching from one DMARD to the other (MTX/LEF group [36.4%] and LEF/MTX [50.0%]). ACR50 rates were higher in the PBO/LEF group (23.8%) than in the other two groups (MTX/LEF 21.2% and LEF/MTX 16.7%).

Safety

Serious adverse events were reported more frequently in the LEF/MTX group (20.0%) than in PBO/LEF (16.1%) or MTX/LEF (11.4%) groups, with only a few considered as related to study drug administration (4.0% LEF/MTX, 3.6% PBO/LEF, and 0% MTX/LEF). Withdrawals due to serious adverse events were infrequent in all treatment groups (4.0% LEF/MTX, 3.6% PBO/LEF, and 0% MTX/LEF). Serious adverse events that occurred in more than one subject receiving leflunomide included coronary artery disorder (2 subjects), bone necrosis (2), and joint disorder (2). Adverse events (both serious and non-serious adverse events) were reported more frequently with leflunomide treatment (100.0% PBO/LEF, 94.3% MTX/LEF) than methotrexate (88.0% LEF/MTX).

Adverse events considered related to study drug administration were more frequent in the LEF groups (78.6% PBO/LEF, 77.1% MTX/LEF) than the methotrexate group (56.0% LEF/MTX). The frequency of withdrawals due to adverse events was higher in the LEF groups (19.6% PBO/LEF, 14.3% MTX/LEF) compared to methotrexate treatment (12.0% LEF/MTX), due to a greater incidence of gastrointestinal disorders. The most common adverse events considered related to leflunomide administration were of gastrointestinal origin and consisted predominantly of diarrhea (26.8% PBO/LEF, 28.6% MTX/LEF, 8.0% LEF/MTX), nausea (16.1% PBO/LEF, 17.1% MTX/LEF, 16.0% LEF/MTX), dyspepsia (10.7% PBO/LEF, 8.6% MTX/LEF, 4.0% LEF/MTX), abdominal pain (8.9% PBO/LEF, 5.7% MTX/LEF, 8.0% LEF/MTX), and LFT abnormalities (7.1% PBO/LEF, 11.4% MTX/LEF, 8.0% LEF/MTX). Other adverse events that appeared related to leflunomide were alopecia (14.3% PBO/LEF, 5.7% MTX/LEF, 4.0% LEF/MTX) and rash (7.1% PBO/LEF, 8.6% MTX/LEF, 0% LEF/MTX).

Of the 15 total subjects with non-serious adverse events leading to withdrawal that were related to treatment, 16.1% were in the PBO/LEF, 11.4% in the MTX/LEF, and 8.0% in the LEF/MTX groups. Of these, the most frequent events (5 subjects, 8.9%) were in the digestive body system, and included LFT abnormalities, vomiting, diarrhea, nausea, and aphthous stomatitis.

The overall incidence of infections was higher in the methotrexate group compared to both LEF groups (64.0% LEF/MTX, 57.1% PBO/LEF, 54.3%

MTX/LEF). The majority of infections as adverse events were respiratory infections, which occurred at a higher rate in the LEF/MTX group (44.0%), compared with the PBO/LEF (33.9%) and MTX/LEF (22.9%) groups.

The most frequently reported adverse events in the cardiovascular system were hypertension (14.3% MTX/LEF, 7.1% PBO/LEF, 0% LEF/MTX), and chest pain (11.4%, 5.4%, and 8.0%, respectively). Hypertension adverse events were related to study drug in 11.4% MTX/LEF and 1.8% PBO/LEF subjects. In the subset of subjects with hypertension as an adverse event, there was a higher incidence of concomitant hypertension at baseline in 2 (3.6%) PBO/LEF, and 4 (11.4%) MTX/LEF subjects, compared to 0% LEF/MTX subjects. New-onset hypertension was reported more frequently in LEF subjects (3.6% PBO/LEF, 2.9% MTX/LEF, 0% LEF/MTX), but the incidence was low. There were no treatment discontinuations due to hypertension.

Potential allergic reactions were reported in 28.6% PBO/LEF, 25.7% MTX/LEF, and 8.0% LEF/MTX subjects. Potential allergic reactions possibly or probably related to leflunomide consisted mainly of rash (7.1% PBO/LEF and 8.6% MTX/LEF, compared to 0% LEF/MTX).

The incidence of RA-related adverse events was evenly distributed among the three treatment groups (25.0% PBO/LEF, 22.9% MTX/LEF, and 24.0% LEF/MTX).

Central nervous system adverse events were reported in 32.1% PBO/LEF, 31.4% MTX/LEF, and 24.0% LEF/MTX subjects. The majority of central nervous system adverse events consisted of headache (21.4% PBO/LEF, 14.3% MTX/LEF, and 24.0% LEF/MTX), paresthesia (10.7%, 11.4%, and 4.0%), and dizziness (7.1%, 5.7%, and 0%). There was one serious adverse event in the MTX/LEF group, consisting of anxiety and paresthesia that was judged unrelated to study drug administration.

Laboratory variables. Leflunomide administration had no effect on sodium, potassium, chloride, bicarbonate, BUN or creatinine. Leflunomide appeared to be associated with a decrease in uric acid, due to the known effect of the drug on the brush border membrane of the proximal renal tubule cells without alterations in renal function or evidence of renal tubular acidosis.

All treatment groups at any visit had similar elevations of SGPT (ALT) and SGOT (AST). Most elevations were mild to moderate and resolved during treatment; moderate elevations of SGPT (ALT) (> 2x ULN to 3x ULN) were highest in the LEF/MTX group (8.0%), compared to the PBO/LEF (3.6%), and MTX/LEF (5.7%) groups. Marked elevations at worst evaluation (>3x ULN) were less frequent and reversed without dose reduction. Liver function abnormalities that were judged related to study drug were similar in all treatment groups, and occurred in 4 (7.1%) PBO/LEF, 4 (11.4%) MTX/LEF, and 2 (8.0%) LEF/MTX subjects. Of these, 1 (1.8%) PBO/LEF, 1 (2.9%) MTX/LEF, and 2 (8.0%) LEF/MTX subjects discontinued from the study. Leflunomide administration did not appear to be associated with any clinically significant adverse effects on SGPT (ALT) and SGOT (AST) in the alternate therapy cohort.

Alkaline phosphatase was mildly elevated (> 1.2 x ULN and < 2 x ULN) in 8.9% PBO/LEF, 5.7% MTX/LEF, and 4.0% LEF/MTX subjects. Moderate elevation (> 2 x ULN to < 3 x ULN) occurred in 2.9% of MTX/LEF subjects. Leflunomide appeared to be associated with a mild, but clinically

insignificant elevation of alkaline phosphatase in a small percentage of patients.

There was no effect on total protein, albumin or total bilirubin and no clinically significant differences in urinalysis results between the three treatment groups.

Leflunomide was associated with an increase in triglycerides, with shifts from normal values at baseline to values above the normal range at endpoint in 20.4% PBO/LEF, 26.9% MTX/LEF, and 0% LEF/MTX subjects. Leflunomide did not appear to be associated with any clinically significant increase in total cholesterol.

There were no adverse effects of leflunomide on other chemistry parameters, such as glucose, creatine kinase (CK), calcium, and phosphorous.

Clinical Variables. There were no clinically significant differences between the three treatment groups in ECG, chest X-ray or physical examination results. Leflunomide had no effect on body temperature or heart rate. Clinically relevant changes in weight from baseline to endpoint occurred in 3.6% PBO/LEF, 8.6% MTX/LEF, and 0% of LEF/MTX subjects. Hence, weight loss in a small number of subjects may be associated with leflunomide.

Subjects in both the leflunomide treated groups in the alternate therapy phase had similar safety profiles, regardless of the treatment during the initial therapy phase.

COMMENTS/ CONCLUSIONS

This clinical trial yielded highly statistically significant results that demonstrated leflunomide was superior to placebo in the treatment of active RA. This was demonstrated by a reduction of signs and symptoms of RA, and the sustaining of clinical and radiographic benefit over 24 months. The retardation of disease progression was maintained between year 1 and year 2 in subjects who continued therapy. Analysis of the functional ability and health-related quality of life measures over 2 years showed the use of leflunomide effected statistically significant improvements over both short- and long-term therapy. In addition, ACR20 response rates for the leflunomide and methotrexate groups were statistically equivalent, though the leflunomide rate was numerically higher at 24 months. There was an earlier response in the LEF group compared to the MTX group, although the results for methotrexate approached those of leflunomide at endpoint. The methotrexate response was highly significantly better than the placebo response. The safety profile of leflunomide in the ITT cohort over 2 years, and the year-2, and alternate therapy cohorts appeared to be acceptable and was generally similar to that of methotrexate. These results provide a safety profile that supports the findings from the one year data.

HWA 486/F/USA/301/RA - LEFLUNOMIDE

HAQ and PET

Site Code | Patient ID Number | Patient Initials FI MI LI

Date Form Completed M D Y | Study Week Baseline ☐ 1 24 ☐ 2 52 ☐ 3 Study Exit ☐ 4

The following questions ask more specifically about the level of difficulty you have when performing specific activities, and the importance of these activities to you. Please mark the place on each line which is closest to the way you feel.

Dressing and Grooming

1. Dressing yourself, including tying shoelaces and doing buttons:

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Frequency of difficulty: 0 (never) — occasionally — frequently — 7 (always)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

2. Are you able to shampoo your hair?

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Frequency of difficulty: 0 (never) — occasionally — frequently — 7 (always)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

Arising

3. Are you able to stand up from a straight chair?

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

4 Are you able to get in and out of bed?

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

HWA 486/F/USA/301/RA - LEFLUNOMIDE

HAQ and PET

Site Code | Patient ID Number | Patient Initials FI MI LI

Date Form Completed M D Y | Study Week Baseline ☐ 1 24 ☐ 2 52 ☐ 3 Study Exit ☐ 4

Eating

5. Are you able to cut your meat?

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

6. Are you able to lift a full cup or glass to your mouth?

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

7. Are you able to open a milk carton?

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

Walking

8 Are you able to walk outdoors on flat ground?

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

9 Are you able to climb up five steps?

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

HWA 486/F/USA/301/RA - LEFLUNOMIDE

HAQ and PET

Site Code | Patient ID Number | Patient Initials FI MI LI

Date Form Completed M D Y | Study Week Baseline ☐1 24 ☐2 52 ☐3 Study Exit ☐4

10. Please check any AIDS OR DEVICES that you usually use for dressing, arising, eating, or walking:

☐1 Cane ☐1 Walker ☐1 Crutches ☐1 Wheelchair
☐1 Devices for dressing (button hook, zipper pull, long-handled shoehorn)
☐1 Built up or special utensil
☐1 Special or built up chair
☐1 Other Specify _____

11. Please check any categories for which you usually need HELP FROM ANOTHER PERSON (check all that apply):

☐1 Dressing and grooming ☐1 Eating
☐1 Arising ☐1 Walking

Hygiene

12. Are you able to wash and dry your body?

Level of difficulty: 0 (without any difficulty) — (some difficulty) — (much difficulty) — 7 (unable to do)

Frequency of difficulty: 0 (never) — (occasionally) — (frequently) — 7 (always)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

13. Are you able to take a tub bath?

Level of difficulty: 0 (without any difficulty) — (some difficulty) — (much difficulty) — 7 (unable to do)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

14 Are you able to get on and off the toilet?

Level of difficulty: 0 (without any difficulty) — (some difficulty) — (much difficulty) — 7 (unable to do)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

Reach

15 Are you able to reach and get down a five pound object (such as a bag of sugar) from just above your head?

Level of difficulty: 0 (without any difficulty) — (some difficulty) — (much difficulty) — 7 (unable to do)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

+ +

HWA 486/F/USA/301/RA - LEFLUNOMIDE

HAQ and PET

Site Code Patient ID Number Patient Initials FI MI LI

Date Form Completed M D Y Study Week Baseline ☐ 1 24 ☐ 2 52 ☐ 3 Study Exit ☐ 4

16. Are you able to bend down to pick up clothing from the floor?

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

Grip

17 Are you able to open car doors?

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

18 Are you able to open jars which have been previously opened?

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

19 Are you able to turn faucets on and off?

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

Activities

20. Are you able to run errands and shop?

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Severity of difficulty: 0 (none) — mild — moderate — 7 severe

Importance of this ability: 0 (not important at all) — 7 (extremely important)

HWA 486/F/USA/301/RA - LEFLUNOMIDE

HAQ and PET

Site Code | Patient ID Number | Patient Initials FI MI LI

Date Form Completed M D Y | Study Week Baseline ☐ 1 24 ☐ 2 52 ☐ 3 Study Exit ☐ 4

21. Are you able to get in and out of a car?

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Severity of difficulty: 0 (none) — mild — moderate — 7 (severe)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

22. Are you able to do chores such as vacuuming or yard work?

Level of difficulty: 0 (without any difficulty) — some difficulty — much difficulty — 7 (unable to do)

Importance of this ability: 0 (not important at all) — 7 (extremely important)

23. Please check any AIDS OR DEVICES that you usually use for self-care, reaching, and gripping.

- ☐ 1 Raised toilet seat
- ☐ 1 Bathtub seat
- ☐ 1 Jar opener (for jars previously opened)
- ☐ 1 Other (specify): _____
- ☐ 1 Bathtub bar
- ☐ 1 Long-handled appliances for reach
- ☐ 1 Long-handled appliances in bathroom

24. Please check any categories for which you usually need HELP FROM ANOTHER PERSON.

- ☐ 1 Hygiene
- ☐ 1 Reach
- ☐ 1 Gripping and opening
- ☐ 1 Errands and chores

25. Please place a mark on the line below that reflects how you feel about your overall health.

Worst Possible Health — Perfect Health

1 2 3 4 5 6 7 8 9 10

HWA 486/F/USA/301/RA - LEFLUNOMIDE

MHAQ

Site Code Patient ID Number Patient Initials FI MI LI

Date of Evaluation M D Y

Visit (Specify study week code)

Study Phase:
☐ 1 Initial
☐ 2 Alternate Therapy

The questions below concern your daily activities. The few minutes you spend answering these questions can provide a more complete picture of how a medical condition may affect your life, adding to information from standard medical tests such as blood tests and X-rays. Please try to answer each question, even if you do not think it is related to you or any condition you may have. Please answer exactly as you think or feel, as there are no right or wrong answers

TO BE COMPLETED BY PATIENT OR STUDY COORDINATOR IF PATIENT UNABLE.

Since your last visit, as a result of your arthritis or any side effects:

1 Have you missed any time at work? Yes ☐ 1 No ☐ 2

2. Were you unable to perform your daily activities? Yes ☐ 1 No ☐ 2

} **If either is Yes, complete form PRU.**

Please check the appropriate box:

AT THIS MOMENT, are you able to:	**Without ANY Difficulty**	**With SOME Difficulty**	**With MUCH Difficulty**	**UNABLE To Do**
3. Dress yourself, including tying shoelaces and doing buttons?	☐ 0	☐ 1	☐ 2	☐ 3
4 Get in and out of bed?	☐ 0	☐ 1	☐ 2	☐ 3
5 Lift a full cup or glass to your mouth?	☐ 0	☐ 1	☐ 2	☐ 3
6. Walk outdoors on flat ground?	☐ 0	☐ 1	☐ 2	☐ 3
7 Wash and dry your entire body?	☐ 0	☐ 1	☐ 2	☐ 3
8 Bend down to pick up clothing from the floor?	☐ 0	☐ 1	☐ 2	☐ 3
9 Turn regular faucets on and off?	☐ 0	☐ 1	☐ 2	☐ 3
10 Get in and out of a car?	☐ 0	☐ 1	☐ 2	☐ 3
11 Duration of morning stiffness				minutes

12. **How Much Pain Have You Had Because Of Your Arthritis IN THE PAST WEEK?**
Mark an X on the scale for how severe your pain has been.

NO PAIN — PAIN AS BAD AS IT COULD BE

13. **Considering All Ways Arthritis Affects You, Mark an X on the scale for how well you are doing.**

VERY WELL — VERY POORLY

HWA 486/F/USA/301/RA - LEFLUNOMIDE

SF-36, MOS Current Health, Work Productivity

Site Code Patient ID Number Patient Initials FI MI LI

Date M D Y

Study Week

Baseline ☐ 1 24 ☐ 2 52 ☐ 3 Study Exit ☐ 4

QUALITY OF LIFE QUESTIONNAIRE

Instructions: This survey asks for your views about your health. This information will help keep track of how you feel and how well you are able to do your usual activities.

Answer every question by marking the answer as indicated. If you are unsure about how to answer a question, please give the best answer you can.

	(Circle One Number)				
	EXCELLENT	VERY GOOD	GOOD	FAIR	POOR
1 In general, would you say your health is:	1	2	3	4	5

	(Circle One Number)				
	MUCH BETTER NOW THAN ONE YEAR AGO	SOMEWHAT BETTER NOW THAN ONE YEAR AGO	ABOUT THE SAME AS ONE YEAR AGO	SOMEWHAT WORSE NOW THAN ONE YEAR AGO	MUCH WORSE NOW THAN ONE YEAR AGO
2 **Compared to one year ago,** how would you rate your health in general **now?**	1	2	3	4	5

ACTIVITIES

3. The following items are about activities you might do during a typical day Does **your health now limit you** in these activities? If so, how much?	(Circle One Number on Each Line) YES LIMITED A LOT	YES LIMITED A LITTLE	NO, NOT LIMITED AT ALL
a) **Vigorous activities**, such as running, lifting heavy objects, participating in strenuous sports	1	2	3
b) **Moderate activities**, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf	1	2	3
c) Lifting or carrying groceries	1	2	3
d) Climbing **several** flights of stairs	1	2	3
e) Climbing **one** flight of stairs	1	2	3
f) Bending, kneeling, or stooping	1	2	3
g) Walking **more than a mile**	1	2	3
h) Walking **several blocks**	1	2	3
i) Walking **one block**	1	2	3
j) Bathing or dressing yourself	1	2	3

HWA 486/F/USA/301/RA - LEFLUNOMIDE

SF-36, MOS Current Health, Work Productivity

Site Code Patient ID Number Patient Initials FI MI LI

Date M D Y

Study Week

Baseline [ ] 1 24 [ ] 2 52 [ ] 3 Study Exit [ ] 4

(Circle One Number on Each Line)

		YES	NO
4	During the **past 4 weeks**, have you had any of the following problems with your work or other regular daily activities **as a result of your physical health**?		
a)	Cut down on the **amount of time** you spent on work or other activities	1	2
b)	**Accomplished less** than you would like	1	2
c)	**Were limited in the kind** of work or other activities	1	2
d)	**Had difficulty** performing the work or other activities (for example, it took extra effort)	1	2

(Circle One Number on Each Line)

		YES	NO
5	During the **past 4 weeks**, have you had any of the following problems with your work or other regular daily activities **as a result of any emotional problems** (such as feeling depressed or anxious)?		
a)	Cut down on the **amount of time** you spent on work or other activities	1	2
b)	**Accomplished less** than you would like	1	2
c)	Didn't do work or other activities as carefully as usual	1	2

(Circle One Number)

	NOT AT ALL	SLIGHTLY	MODERATELY	QUITE A BIT	EXTREMELY
6 During the **past 4 weeks**, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbors, or groups?	1	2	3	4	5

(Circle One Number)

	NONE	VERY MILD	MILD	MODERATE	SEVERE	VERY SEVERE
7. How much bodily pain have you had during the **past 4 weeks**?	1	2	3	4	5	6

(Circle One Number)

	NOT AT ALL	A LITTLE BIT	MODERATELY	QUITE A BIT	EXTREMELY
8 During the **past 4 weeks**, how much did pain interfere with your normal work (including both work outside the house and housework)?	1	2	3	4	5

HWA 486/F/USA/301/RA - LEFLUNOMIDE

SF-36, MOS Current Health, Work Productivity

Site Code Patient ID Number Patient Initials FI MI LI

Date M D Y

Study Week

Baseline ☐ 1 24 ☐ 2 52 ☐ 3 Study Exit ☐ 4

(Circle One Number on Each Line)

9. These questions are about how you feel and how things have been with you during the past 4 weeks. For each questions, please give the one answer that comes closest to the way you have been feeling. How much of the time during the past 4 weeks...	ALL OF THE TIME	MOST OF THE TIME	A GOOD BIT OF THE TIME	SOME OF THE TIME	A LITTLE OF THE TIME	NONE OF THE TIME
a) Did you feel full of pep?	1	2	3	4	5	6
b) Have you been a very nervous person?	1	2	3	4	5	6
c) Have you felt so down in the dumps that nothing could cheer you up?	1	2	3	4	5	6
d) Have you felt calm and peaceful?	1	2	3	4	5	6
e) Did you have a lot of energy?	1	2	3	4	5	6
f) Have you felt downhearted and blue?	1	2	3	4	5	6
g) Did you feel worn out?	1	2	3	4	5	6
h) Have you been a happy person?	1	2	3	4	5	6
i) Did you feel tired?	1	2	3	4	5	6

10. During the past 4 weeks	ALL OF THE TIME	MOST OF THE TIME	SOME OF THE TIME	A LITTLE OF THE TIME	NONE OF THE TIME
a) How much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)	1	2	3	4	5

HWA 486/F/USA/301/RA - LEFLUNOMIDE

SF-36, MOS Current Health, Work Productivity

Site Code Patient ID Number Patient Initials FI MI LI

Date M D Y

Study Week

Baseline ☐ 24 ☐ 52 ☐ Study Exit ☐

(Circle One Number on Each Line)

11 How TRUE or FALSE is each of the following statements for you?	DEFINITELY TRUE	MOSTLY TRUE	DON'T KNOW	MOSTLY FALSE	DEFINITELY FALSE
a) I seem to get sick a little easier than other people	1	2	3	4	5
b) I am as healthy as anybody I know	1	2	3	4	5
c) I expect my health to get worse	1	2	3	4	5
d) My health is excellent	1	2	3	4	5
e) I am somewhat ill	1	2	3	4	5
f) I am in poor health	1	2	3	4	5
g) I have been feeling bad lately	1	2	3	4	5
h) I feel as good as I ever have	1	2	3	4	5

12. How happy, satisfied, or pleased have you been with your health during the **past 4 weeks**?

EXTREMELY HAPPY, COULD NOT HAVE BEEN MORE SATISFIED OR PLEASED	VERY HAPPY MOST OF THE TIME	GENERALLY SATISFIED, PLEASED	SOMETIMES FAIRLY SATISFIED, SOMETIMES FAIRLY UNHAPPY	GENERALLY DISSATSIFIED, UNHAPPY	VERY DISSATISFIED, UNHAPPY MOST OF THE TIME
1	2	3	4	5	6

These following questions refer to WORK

(Circle One Number)

13 During the **past four weeks**,..	PAID WORK	HOUSE WORK	SCHOOL WORK	UNEM-PLOYED	DISABLED	RETIRED
a) What has been your main form of work?	1	2	3	4	5	6

If you answered unemployed, disabled, or retired, please go to end of questionnaire

(Circle One Number on Each Line)

14 During the **past four weeks**, .	ALL DAYS	MOST DAYS	SOME DAYS	FEW DAYS	NO DAYS
a) How often were you unable to do any paid work, house work, or school work?	1	2	3	4	5
b) On the days that you did work, how often did you have to work a shorter day?	1	2	3	4	5
c) On the days that you did work, how often were you unable to do your work as carefully and accurately as you would like?	1	2	3	4	5
d) One the days that you did work, how often did you have to change the way your paid work, house work or school work is usually done?	1	2	3	4	5

+ +

HWA 486/F/USA/301/RA - LEFLUNOMIDE

Work Problems

Site Code Patient ID Number Patient Initials FI MI LI

Date M D Y

Study Week

Baseline ☐ 24 ☐ 52 ☐ Study Exit ☐

(Circle One Number on Each Line)

15. During the **past four weeks**, how much difficulty have you had doing the following work activities because of any ongoing health problems or health concerns?	NONE	A SLIGHT AMOUNT	SOME	QUITE A BIT	A GREAT DEAL	NOT DONE, CAN'T DO IT	NOT PART OF JOB
a) Keeping up with required standards of personal appearance, dress, personal safety and hygiene?	1	2	3	4	5	6	7
b) Maintaining required attitudes toward your work? (includes morale, motivation, commitment)	1	2	3	4	5	6	7
c) Doing required commuting or local and long-distance traveling?	1	2	3	4	5	6	7
d) Walking or moving around your usual work area or building?	1	2	3	4	5	6	7
e) Doing things that require you to concentrate, remember, make decisions, solve problems, or make judgments? (not being too distracted by health problems or concerns)	1	2	3	4	5	6	7
f) Doing things that require you to use or move all or part of your body in your work? (staying in one position, staying in awkward or unusual positions, repeating motions, moving or lifting, exerting yourself)	1	2	3	4	5	6	7
g) Using and controlling small or light-weight devices, tools, machines or equipment? (includes personal computers).	1	2	3	4	5	6	7
h) Using and controlling LARGE or HEAVY devices, tools, machines or equipment?	1	2	3	4	5	6	7
i) Arranging, moving or otherwise physically handling materials used in your work? (objects, animals or people)	1	2	3	4	5	6	7
j) Doing your work well and on time? (maintaining required quality, consistency, workload, work schedules, deadlines, pacing, work-flow, employer or supervisor satisfaction or customer/client satisfaction)	1	2	3	4	5	6	7

THANK YOU FOR ANSWERING THIS QUESTIONNAIRE!

POST-MARKETING COHORT STUDY OF LEFLUNOMIDE AND OTHER DMARDS

A COMPARATIVE RISK ANALYSIS

STUDY REPORT

by

Global Epidemiology

Aventis Pharmaceuticals Inc.

13 January 2003

CONFIDENTIAL

POST-MARKETING COHORT STUDY OF LEFLUNOMIDE AND OTHER DMARDS

A COMPARATIVE RISK ANALYSIS

1. BACKGROUND

A post-marketing, retrospective cohort study comparing the rate of adverse events (AEs) amongst leflunomide users to patients taking Disease Modifying Anti-Rheumatic Drugs (DMARDs) (eg, gold salts, azathioprine, hydroxychloroquine, D-penicillamine, and sulphasalazine), methotrexate, NSAIDs, and the cox-2 inhibitors was undertaken to address several issues. Leflunomide, approved by the FDA in September 1998, was the first new DMARD introduced in a decade, and is indicated for adults with active rheumatoid arthritis (RA) to reduce signs and symptoms and to retard structural damage as evidenced by X-ray erosions and joint space narrowing. Spontaneous reports to drug regulators as well clinical cases described in the literature warrant a formal investigation of these potential signals. Specifically, this study focuses on the assessment of serious hepatic, dermatologic, hematologic, infectious and other adverse events.

2. PRINCIPAL AIMS

The principal aim of this cohort study was to:

- determine the incidence rates of serious hepatic (e.g., liver necrosis, hepatitis, acute liver failure), dermatologic (e.g., Stevens-Johnson Syndrome, Toxic Epidermal Necrolysis), hematologic (e.g., aplastic anemia), hypertension, and other adverse events among leflunomide users and compare them to rates in users of other DMARDs, NSAIDs, cox-2 inhibitors, and methotrexate.

3. STUDY METHODS AND DATA SOURCE

3.1 STUDY DESIGN

The design is a dynamic retrospective cohort study. The advantage of this design is its flexibility to identify a large number of patient cohorts, defined by diagnosis and medication exposure, and follow them through their course of therapy to estimate directly the strength of the association between exposure and outcome (e.g., severe hepatic and other events) over time.

3.2 DATA SOURCE

The data source selected for this study is the Aetna-US Healthcare claims database, a repository of health information on 6,470,000 covered lives, with linkage to medical, pharmacy, and laboratory data. Aetna has a long history of clinical database research (1, 2) and an established infrastructure for working with databases, as well as the internal capability to abstract data from medical records. Insurance claims databases are commonly used in pharmacoepidemiology research, as they afford the investigator large numbers of subjects (often the largest available) as well as data on medical services, pharmacy services (including date of dispensing, drug name and dosage, and duration of prescription), and enrollment time of members (which allows calculation of person time at risk).

3.3 STUDY PERIOD

Leflunomide was launched in September 1998 in the US. All leflunomide users were identified in the dataset and similarly all comparator subjects were identified from September 1998 through December 2000.

3.4 DEFINITION OF STUDY COHORTS

Several cohorts were defined for this study. As a reference cohort, leflunomide monotherapy and leflunomide + methotrexate patients were used, given that they are of most interest. Comparison cohorts include methotrexate, DMARD, and non-DMARD patients (NSAID and cox-2 users) as well as leflunomide plus methotrexate and leflunomide plus other DMARD patients. Comparisons were limited to monotherapy and two-drug therapy cohorts. The design is dynamic (or variable) meaning that persons may contribute exposure time to any number of different cohorts. Time windows of exposure are defined below.

3.5 DEFINITION OF EXPOSURE PERIOD BASED ON PHARMACY FILES

Classification of each cohort member's person-time began with the first prescription of any of the exposures of interest (leflunomide, methotrexate, DMARDs, and non-DMARDs) and proceeded through the end of the last prescription. To take account of leflunomide's relatively long half-life (estimated to be two weeks), the washout period was set at 60 days. Thus, the person-time for a single prescription includes each day from the date of the first prescription through either the last day for that prescription (plus 60 days) or the first day of the next prescription, whichever comes first. Similarly for the other exposures of interest, the equivalent of five half-lives were added to the end of the prescription period in order to calculate person-time exposure (see Appendix C for the list of half-lives).

In the case where a person has overlapping prescriptions for different rheumatoid arthritis medications, his or her person-time was apportioned to the appropriate combination exposure category for that period of overlap time. This was easily accomplished with the Aetna database, which records the following: days supply, units dispensed, strength, and date of dispensing.

For any dispensing for which the days supplied are missing or zero, the median days supplied for users of that medication was employed.

Person-time at risk was aggregated into the different time windows according to leflunomide or other DMARD use and continued until the earliest 1) confirmed event of interest; 2) end of washout for a given medication, 3) date of last enrollment; 4) death; or 5) end of the study period. Again, this is a dynamic cohort in which a subject may contribute person-time to more than one cohort.

Exposure period at risk will end with the first of these events:

- End of the study period
- Termination of enrollment in the health plan. Because small lapses in enrollment are not uncommon due to administrative procedures, gaps in enrollment of up to 31 days are permitted.
- Specific clinical outcomes of interest
- Death

The assumption is that the outcome of interest is acute, in that it has a close temporal link to the exposure. To study this type of drug effect it is necessary to track closely how drug use changes over time, because the drug-induced risk begins when the drug is started. The half-life of leflunomide is approximately two weeks and therefore the risk period for leflunomide patients was defined as the time on leflunomide therapy plus 60 days (approximately 5 times the metabolic half-life) after the last dose. Similarly, the methotrexate risk period was defined as the time on methotrexate therapy plus 125 days. The half-lives of other DMARDs varied from 0.1 to 27 days, and the exposure tails were calculated accordingly.

In quantifying exposure to the study medications, attempts were made to track changes that occur with time. Because it is never definitively known how patients are taking their medication or when they actually stop, only proxies for this exposure, as per instructions of prescriptions filled, can be created.

3.6 INCLUSION CRITERIA INTO THE COHORT

- Because several of the comparator medications have indications other than rheumatoid arthritis, inclusion in the study population for non-leflunomide patients was defined by a combination of diagnosis and treatment. Therefore, all leflunomide patients were included in the study (because the drug is indicated only for RA) while comparator drug users (ie, exposures to the other medications listed in Appendix A) required a rheumatoid arthritis diagnosis within 90 days (before or after) the prescription date for the medication, to assure limiting the analyses to patients with rheumatoid arthritis.
- Sex and date of birth are known.
- Eighteen years of age or older on t_0, the time of entry into the cohort.

3.7 EXCLUSION CRITERIA

Amongst the non-leflunomide users, patients were excluded if they experienced one of the hepatic events of interest (see Table 1) in the 90-day period prior to potential entry into the cohort. The standard practice of excluding patients with the events of interest is critical to the chosen design--participants must be at risk of developing the outcomes of interest. All leflunomide users were included in the study, those with and without a history of hepatic disease (this issue is addressed later in the report).

3.8 PRIMARY ENDPOINTS

Any inpatient or outpatient encounter that reflects the key events of interest will be considered a potential case. It is recognized that these codes in and of themselves do not connote a medication-induced event but they do allow the casting of a wider net and do not exclude potential cases that may be of interest. ICD codes for serious events are specified in the protocol (see Appendix B). As an example, Table 1 lists several codes for hepatic events.

Other primary endpoints include hematologic (acquired pancytopenia, aplastic anemia); severe skin reactions (Stevens-Johnson syndrome, toxic epidermal necrolysis); hypertension; vasculitis and hemolytic anemia; pneumonitis; acute pancreatitis; GI bleeding; and upper respiratory tract infections and bronchitis.

4. ANALYTIC PLAN

Simple descriptive characteristics of the cohort have been generated, in addition to total subjects, person-time, mean length of exposure time, and number of events. Incident rates have been calculated to compare events between leflunomide monotherapy patients and a series of comparator patients (eg, methotrexate monotherapy and various combinations of drugs), along with 95% confidence intervals to facilitate evaluation of product differences.

Adjustment for potential confounders was performed by putting age, sex, and comorbidity into a Poisson regression model. The Poisson assumption was chosen for the modeling strategy because it presumes that the number of outcomes of interest are small compared to the total cohort size and are statistically independent events. This assumption holds even if the same individuals contribute person-time to more than one stratum. Poisson regression theory presumes that the rarity of the outcome events in any one time frame, whilst removing those individuals who experienced those events from risk, nonetheless has little effect on the probability of a specified number of events in the next time frame(3-6).

Adjustment for comorbidities utilized the scoring of the Charlson Index (7). This is a weighted index used to classify comorbidities, taking into account their number and severity.

5. CASE VALIDATION

An important element of this study is case validation. As originally described in the study protocol, the validation process will include primary data abstraction from the source medical records according to an established procedure. The validation effort will be directed only to hepatic events; the total number of AEs observed in the cohort is 16 000, which is beyond the scope of any validation endeavor. The final sample of hepatic events to be validated will be selected from the total of 651 such events, as follows: 100% of the liver necroses; 20% of the biliary cirrhosis cases; 25% of hepatic coma cases; 25% of noninfectious toxic hepatitis cases; 10% of the non-alcoholic liver cirrhosis cases; 12% of unspecified chronic liver disease; 10% of the other specified liver disorders; 10% of the unspecified liver disorder; 12% of the jaundice cases; .and 10% of the elevated liver enzyme cases.

Detailed data on the severe hepatic events were obtained. Approximately 61% of these events were noninfectious, toxic hepatitis; 4% were biliary cirrhosis, 4% acute necrosis of the liver, 2% hepatic coma, and 28% were 'orphan events', ie, those AEs not associated with any drug exposure because they did not occur within the defined exposure windows.

The formal protocol will be as follows:

- All patients in the RA cohort with a diagnosis code indicating a hepatic event of interest will be identified. A letter will be sent from the US Quality Assurance (USQA) department at Aetna to the site of care requesting copies of both the office/hospital notes at and around the time of the coded event of interest in addition to any of the following laboratory tests that may have been undertaken: liver biopsy, ultrasound, CT/MRI scan, serum chemistries including liver enzymes, bilirubin and hepatitis titres.
- The office or hospital will receive a financial incentive to respond.
- All responses will be de-identified by USQA.
- All responses will be collected centrally by USQA.
- A trained nurse-abstractor will review all returned material and complete as much as possible the attached form regarding the patient's pre- and post-diagnosis status as outlined in the abstract form in Appendix D (ref: P*h*RMA/FDA/AASLD Drug-Induced Hepatotoxicity White Paper: Postmarketing Considerations. November 2000)

The validation by record review is currently in progress and results will be presented later when the data are available to us.

6. RESULTS: THE STUDY COHORT

During the study period, detailed below, a total of 40 954 RA patients were identified. These patients represent approximately 83 143 person-years exposure to one or two drug therapy, including 11 130 person-years exposure to leflunomide. The demographics of the patients are presented in Table 2. The distribution of total person-time is shown in Figure 1. There were no major disparities between or amongst cohorts. For example, males aged 18-30 contributed approximately 1% of the person time, averaged across the five monotherapy groups, females aged 18-30 contributed 2-4% of the person time.

The female:male ratio is 2.7, confirming estimates from the literature.

The cohort itself is drawn from the larger Aetna-US Healthcare population of 6 470 000 persons. Thus, the rate of RA in the Aetna database is 40 594/6 470 000 or 0.63% which is, again, in accord with estimates from the literature.

An assessment of the different cohorts' comorbidities was undertaken to determine their comparability. Due to the time-on-drug dependent nature of the cohorts, such a comorbidity analysis was limited to the monotherapy groups.

The comorbidity analysis examined 72 different conditions prior to the index date, ie, the date at which a person was included in the cohort (Figure 2). There appeared to be a lower number of comorbidities at

the index date amongst the leflunomide (mean=1.62) and methotrexate (mean=1.8) monotherapy compared to the DMARD group (mean=2.7). Figure 3 shows the number of comorbidities amongst leflunomide and non-leflunomide users who experienced an event of interest; no major differences are seen between the groups. The comorbidities are potential confounders and, as such, will be included in the Poisson regression model. The cumulative person-year exposure, mean exposure time, and the number of patients in the database on the therapies of interest are displayed in Table 3.

DMARD + methotrexate users were very common (ie, account for >10% of total person-year exposure). In terms of mean exposure time, use of any DMARD monotherapy and the combination DMARD + MTX were the longest, with a mean of about 2.1 years of exposure. Leflunomide monotherapy had a mean exposure of 1.6 years.

The primary aim of this study was to determine the rate of serious outcomes associated with the use of certain drugs commonly used to treat rheumatoid arthritis. Incidence rates were calculated using varying 'exposure tails' to avoid bias (8).

The incident rates of various outcomes of interest, in total and separately, are presented in Table 4-15. When reviewing the rates, it is critical to remember that in this dynamic cohort, patients can move from one therapeutic category to another (by their prescribers), depending on efficacy and adverse events experienced.

7. RESULTS: ADVERSE EVENT RATES

7.1 ANY ENDPOINT

Leflunomide had the lowest rates of any endpoint examined in this study compared to the other monotherapies (Table 4). This rate was statistically significantly lower than MTX and DMARD. There were 94 events per 1000 person-years observed (adjusted for age, sex, and comorbidities), compared to almost 144 events per 1000 PY amongst DMARD users and 145 events per 1000 PY amongst methotrexate users.

Leflunomide + MTX had an any endpoint rate of 43 per 1000 PY, a rate significantly lower than the DMARD + methotrexate group (70 events per 1000 PY) (Table 4), and marginally significant compared to leflunomide + DMARD (59 events per 1000 PY). The non-DMARD group had the highest rate, 382 events per 1000 PY.

7.2 HEPATIC EVENTS

A total of 644 hepatic events were observed in this cohort. Leflunomide had the lowest rate in the monotherapy groups (4 per 1000 PY; Table 5), although this rate was not significantly different from the other monotherapies. The rate for leflunomide + methotrexate was not significantly different than the comparator two-drug therapies (Table5).

7.3 INDIVIDUAL HEPATIC EVENTS

A detailed presentation of the individual severe and other hepatic events is shown in Table 6. The total numbers of severe hepatic events and other hepatic events tended to be lower in the leflunomide than in other monotherapy groups. The numbers of these events were very low, almost all in the single digits, and therefore differences were not statistically significant. For example, only one case of liver necrosis was observed in leflunomide monotherapy patients, two in methotrexate patients, two in the no-DMARD group, and 7 in the DMARD group. No cases of liver necrosis were observed in the combination therapy groups. Non-infectious hepatitis was the most commonly observed liver event; small numbers again precluded clear trends. No cases of cirrhosis or biliary cirrhosis were seen in the leflunomide patients (monotherapy or combination); three cases of cirrhosis were observed in the methotrexate monotherapy group.

7.4 HEMATOLOGIC EVENTS

Results for hematologic events are shown in Table 8. There were 105 events in the cohort, 19 of which were orphan events (occurring in the apparent absence of drug exposure). Three cases were seen in the leflunomide group versus eight in the methotrexate group. No cases were observed in the leflunomide + methotrexate group.

7.5 SEVERE SKIN REACTIONS

Again, this was an exceedingly rare event: only 32 cases occurred in this study, 5 of which were orphan. No events were seen amongst either leflunomide monotherapy or leflunomide + methotrexate users, and six cases observed in the methotrexate group and 16 in the DMARD group (Table 9).

7.6 HYPERTENSION

Hypertension, a relatively common condition, was relatively common in this study. Lelfunomide had an adjusted rate of 33 per 1000 PY, statistically significantly lower than methotrexate and DMARD monotherapy comparators (the DMARD rate was 48 per 1000 PY and the methotrexate rate was 51 per 1000 PY) (Table 10). Leflunomide + methotrexate also had a relatively low rate, 13 cases per 1000 PY. The non-DMARD group had a rate of 158 per 1000 PY.

7.7 PNEUMONITIS

There were 1038 events in the monotherapy group. The monotherapy rates were comparable: leflunomide, 11 per 1000 PY; methotrexate, 13 per 1000 PY, and DMARD 16 per 1000 PY (Table 12). The combination therapy rates were also comparable.

7.8 PANCREATITIS

The rates of pancreatitis amongst monotherapy and combination therapy patients were similar, ranging from 1.2 to 2.4 per 1000 PY. The non-DMARD group again had the highest rate, 3.6 per 1000 PY.

7.9 RESPIRATORY EVENTS

Respiratory event rates were relatively high, but the lowest was found in the leflunomide group (20 cases per 1000 PY) (Table 15). This rate was statistically significantly lower than the DMARD and methotrexate monotherapy rates, which were about 38 per 1000 PY for both exposures. The combination therapy rates were similar, ranging from 12 to 19 per 1000 PY.

8. DISCUSSION

8.1 EPIDEMIOLOGY OF RHEUMATOID ARTHRITIS

It is critical to understand the epidemiology of rheumatoid arthritis (RA) in order to assess how much disease is occurring, how the disease incidence varies within a population, what the relative disease burden is, and the impact on morbidity and mortality by therapeutic interventions. Knowing the epidemiology of this condition, including its treatments and their associated adverse events, and comorbidities will help not only in the assessment of leflunomide but also in the determination of the validity of the cohort under study, particularly the generalizability of the results to the larger population of RA patients.

Prevalence of RA is easier to ascertain than incidence. The incidence of RA in the population is difficult to determine for a variety of reasons, and has been likened to a moving target(9). The difficulties arise from a lack of consensus on, and changing criteria for, the diagnosis of RA, the often lengthy time between onset of symptoms and contact with health professionals, the natural history of RA, and the relative rarity of the condition, necessitating study in very large populations (and the correlative prohibitive cost of setting up and maintaining population-based disease registries(10)). With these challenges in mind, the prevalence of RA is presented in Table 16.

The table, which is not a complete evaluation from the literature, shows a higher prevalence of RA amongst women (from 2 to almost 4 times the prevalence in men), and prevalence point estimates from 0.29-1.6% amongst women and 0.09-0.66% in men. Total prevalence ranges from 0.34-2.26%, although prevalence exceeding 5% has been reported in certain native American peoples (11-13) .

Table 17 displays incident data on rheumatoid arthritis.

Estimates range from 4.7-49.7 per 100 000 per year amongst men, and from 12.7-98.1 per 100 000 per year amongst women, approximately 8-10-fold differences, which reflect the difficulties underlying determination of incidence. The ratio between women and men ranges from 1.7-2.7; most ratios being approximately 2.7.

8.2 GENERAL DISCUSSION

The discussion will be limited to consideration of the adverse events noted in the monotherapy and two-drug therapy cohorts. Comparisons will be made using leflunomide monotherapy and leflunomide + methotrexate therapy as reference groups.

As noted and defined in the Results section, many of the outcomes of interest were broadly classified. For example, Respiratory Tract Infections included laryngopharyngitis, acute bronchitis, influenza, bronchitis, and other respiratory infections. This conservative approach guards against the formation of diagnostic categories that may be too specific. An exception to this was the detailed analysis of the hepatic events.

Amongst the monotherapy group, leflunomide had the lowest rate of hepatic events; the difference between this rate and those of comparators, however, was not statistically significant. The rate of hepatic events amongst the leflunomide + methotrexate group was also significantly lower than the rates of the comparator two-drug combinations.

Because hepatic events have been viewed with especial interest by regulators, a more detailed analysis of individual events was undertaken. As often happens with analyses of this sort, one ends up with small numbers of events, which can preclude meaningful scrutiny. Such was the case here, for the most part: no events of biliary cirrhosis, hepatic coma, cirrhosis, or unspecified chronic liver disease,were observed in the leflunomide monotherapy cohort. Once all the hepatic events were combined, the leflunomide exposure group had significantly lower rates of AEs than the DMARD cohort ($p = 0.03$).

The methotrexate exposure group had the highest rates of hepatic disease in this study (after the no-DMARD group). Methotrexate is known to be associated with liver damage; there is little direct evidence to support contentions that DMARDs are associated with low risks of adverse events. The vast majority of studies of DMARD toxicity are short-term and are usually of relatively small numbers of selected patients monitored under clinical trial protocols. Further, the results of such studies are often contradicted by long term results from clinical practice which depend on unselected populations of patients (14, 15). Hepatotoxicity is not uncommon amongst users of these drugs, although it is difficult to find consensus in the literature on incidence, partly because there is no agreed upon endpoint and partly because study populations are, for the most part, not comparable (although that is not the case in the present study). A discussion of NSAID hepatotoxicity follows.

NSAID hepatotoxicity was established in 1923 with cinchopen, an anti-rheumatic drug, which has long since been withdrawn from the market(16). Benoxaprofen (Coxigon)(17) and ibufenac(16) were similarly withdrawn for their role in causing hepatic injury (fatal cholestatic jaundice, in the case of the former). One review concludes that because of the potential hepatotoxicity of NSAIDs, liver enzymes should be monitored every 1-2 weeks for eight weeks when therapy is initiated(18). Another review offers the judgment that the incidence of elevated liver enzymes during NSAID use is much higher than the incidence of clinically significant hepatotoxicity and thus, if the indication for NSAID is maintained, there is little risk involved in continuing therapy, unless there are other indices of hepatic injury(19).

Several epidemiologic studies of NSAID-induced hepatotoxicity have been performed. A Danish study reviewed 1100 reports on suspected drug-induced liver injury from 1978 through 1987 and found that about 9% of all hepatic reactions were due to NSAIDs (20). A US-based study using Medicaid billing data examined hospital admissions for acute hepatitis and found an odds ratio of 1.4 (95% confidence interval 0.6, 3.1) for NSAID use (21). Medicaid data, however, are well known for their lack of generalizability, lack of access to confounding variables in the database, and lack of data validation. In this particular study only half the requested records were made available for review, and after exclusions, only 107 cases were analyzed. Another epidemiologic study of hospitalized acute liver injury found a rate of 9 cases per 100 000 person-years amongst NSAID users (22) and another hospital-based study found only three persons out of 102 644 NSAID users to have required admission for acute liver injury (23).

A study combining over 625 000 outpatients and hospitalized patients (from the GPRD database) calculated an incidence of 9 cases of acute liver injury per 100 000 person-years exposure to NSAIDs (24). Interestingly, this study used logistic regression to assess the relative risk for acute liver injury amongst RA patients and found that it was almost 11 times the risk for osteoarthritis patients, the reference group. The risk was also 11 times that of patients with other chronic conditions and over four times the risk amongst patients with 'acute conditions.'

In a recent review of quantitative studies of liver injury amongst NSAID users(25), Walker makes several relevant points about some of the studies discussed above. Reporting is generally incomplete, even amongst hospitalized patients, to the point where it is virtually impossible to assess severity of disease, in this case, the extent of the liver injury. Detection bias is likely if mild elevations of liver enzymes were the outcome of interest, as the occurrence of such elevations would be in direct proportion to the level of surveillance, which is low in the case of NSAIDs (26). Studies of hospitalized patients are hostage to temporal changes in what constitutes a case worth admitting, as well as to barriers to hospitalization, so that one may miss clinically significant cases. Several other noteworthy reviews on this topic are available(27, 28).

The cox-2 inhibitors, currently enormously popular drugs, were launched in 1999 to great acclaim, notably that they were as effective as traditional NSAIDs yet had improved safety profiles. This early acclamation appears to have been premature, as growing numbers of reports are available attesting to similar and more serious adverse events associated with cox-2 use. For example, five cases of hepatoxicity (two of which resulted in death from fulminant hepatic failure) due to a cox-2 were reported in 1998 (29). Another case was reported the following year (30). More recently, reports have surfaced about cox-2 associated acute hepatocellular and cholestatic livery injury (31, 32). There are general reservations about the safety of cox-2 inhibitors (33-35) as well as specific concerns regarding nephrotoxicity (36-38), gastric toxicity (39), thrombotic events (40), and, perhaps most important of all for rheumatoid arthritis patients, the impact of cox-2 inhibition on bone resorption and formation (41). Two sentences from a recent summary should conclude: For painful exacerbations of ostearthritis or rheumatoid arthritis, the

moderate symptomatic effect of celecoxib is no different from that of the other nonsteroidal antiinflammatory drugs with which it has been compared. Furthermore, there is no firm evidence that its safety profile is any more favourable (42).

Studies of the hepatotoxicity of methotrexate and other DMARDs have been conducted. A recent study comparing RA patients with and without viral hepatitis, and in which baseline ductal and parenchymal enzymes were checked prior to commencement of DMARD therapy, found that amongst the RA-only patients, 20% developed abnormal enzyme levels (> two times the upper limit of normal). Although the numbers were small, 25% of azathioprine patients experienced abnormal elevations of ALT, as did 17% of gold patients, 25% of methotrexate patients, and 17% of sulphasalazine patients(43). Combining the different drug combination therapies, about 20% of RA patients experienced abnormal increases in liver enzymes.

In the largest observational study of the adverse effects of DMARDs in the UK, investigators relying upon computerized records of DMARD use and adverse events in an RA clinic found that the rate of liver function abnormality (not defined in the study) was 0.63 per 100 person-years for sulphasalazine, 2.50 per 100 PY for methotrexate, and 2.67 per 100 PY for azathioprine(44). These results are in agreement with those from another large study of long-term results from DMARD therapy, which found the incident rate of hepatic abnormalities (including elevated liver tests, but otherwise not defined) to be 1 per 100 PY amongst methotrexate users(45).

A study in Switzerland of liver function amongst RA patients prior to receiving methotrexate therapy found large increases in values outside the reference range for AST and ALT levels, from 0.8% to 7.6% of patients (AST) and from 5.1% to 10.8% (ALT) (46). The study also documented a continuous quantitative decline in liver function over time.

A study of the toxicity of DMARDs in 2500 RA patients showed that hepatotoxicity occurred amongst methotrexate users with an incidence of 47 events per 1000 person-years(47).

Liver enzymes tests, specifically, aspartate aminotransferase (AST) and alanine aminotransferase (ALT), are plagued with concerns about, amongst others, sensitivity and specificity. For example, how does one assess the importance of elevated liver enzyme tests in asymptomatic patients, how are biochemical reference ranges established, how are the results of such tests used in diagnosing or screening patients who may develop liver disease, and can baseline rates of 'abnormal' be constructed for the general population as well as for selected patient populations? These questions have been asked for decades(48, 49) and the answers remain elusive. Clearly, in the assessment of adverse drug reactions it is critical to establish the values of a normal range(50). Elevated liver enzymes are found in up to 5% of asymptomatic patients(51-53) and in a recent study in the US, a rate of 200 new cases of elevated liver enzymes per 100

000 PY was found in a general population(54). An older study of incident hepatic disorders in a managed care population found a rate of 77 cases per 100 000 PY(55).

It is also imperative to remember that the natural history of rheumatoid arthritis may include liver injury in the absence of therapeutic intervention. Studies of the natural history of this potentially severe condition are, for the most part, not available independently of treatment(56, 57). Although work in this area is fairly old, there is some evidence that RA patients develop biochemical evidence of hepatocellular dysfunction and histologic liver abnormalities(58, 59). A recent autopsy study found fibrosis in 11% of cases reviewed; diffuse fibrosis with no identifiable cause was found in 8.2%(60). Studies in Scotland showed about 13% of RA patients to have definable liver disease(61, 62).

The leflunomide exposure group experienced an adjusted incidence rate of 4 cases of hepatic events per 1000 person-years, as did the DMARD group, compared to 7 events per 1000 PY amongst methotexate users. These rates were not significantly different from one another.

The hematologic events examined in this study (aplastic anemia and pancytopenia) are rare events whose epidemiology has begun to be quantified in the general population but is virtually unknown amongst rheumatoid arthritis patients(63).

In Malaysia, an aplastic anemia (AA) incidence of 4.8 per million person-years was estimated(64); in Thailand, the AA incidence was 3.9 cases per million in Bangkok and the incidence of agranulocytosis was 0.8 per million PY(65, 66). The incidence of toxic agranulocytosis, ie, unrelated to radiation, anticancer drugs, or known industrial toxics, was 8.4 cases per million PY in Buenos Aires(67). An early study, relying on Medicaid billing data and excluding cancer patients and patients receiving cytotoxic and immunosuppresive drugs, calculated an incidence of 7.2 per million PY(68). A recent study utilizing the Saskatchewan database estimated the incidence of AA and agranulocytosis to be 2.7 and 3.0 per million PY(69). A large French case-control study of aplastic anemia found an odds ratio of 6.8 for patients with rheumatoid arthritis, and a related odds ratio of 4.9 for previous use of either gold or penicillamine(70). The association of gold with aplastic anemia amongst RA patients is has been recognized in the past(71). A fatal case of agranulocytosis with sulfasalazine use in RA has been reported(72); this particular adverse event has also been noted previously(73).

Severe leukopenia was observed in eight RA patients in a Canadian study amongst 144 users of low dose methotrexate(74). Reversible leukopenia following sulfasalazine use has been observed in the UK(75) and Germany(76). A large UK study found the rates of netropenia to be 1.49 per 100 PY for sulfasalazine, 1.71 for methotrexate, 0.85 for penicillamine, and 2.14 for azathioprine(44). This study also found rates of 'low platelets' (not otherwise defined) of 0.57 per 100 PY for sulfasalazine, 0.93 for methotrexate, 6.36 for penicillamine, and 2.14 for azathioprine. The present study--for all drugs--has lower rates of hematologic events compared to the rates available in the literature, although this study only captures clinical

diagnoses of aplastic anemia (including pancytopenia) and patients with laboratory abnormalities may not be captured.

There are few data on the epidemiology of severe skin reactions amongst RA patients. Most of the literature concerns case reports in single patients(77-80). In the population at large, the incidence of Stevens Johnson Syndrome (SJS) varies from 1.2 to 6 cases per million person-years and that of Toxic Epidermal Necrolysis (TEN) from 0.4 to1.2 cases per million person-years(81-83). Thirty-two cases were observed in the present cohort (five of which were orphan), resulting in a incidence of 340 per million person-years, about 50 times the expected rate. There is no specific ICD code for SJS/TEN; the code used in this study included SJS, TEN, and erythema multiformae (EM). It is likely that many of the events captured for this particular event were of the latter variety, ie, EM, and were likely mild. Thus, it is not surprising to see a higher than expected event rate. No cases were seen in leflunomide monotherapy users and only one in any of the two-drug leflunomide combinations (leflunomide + DMARD; 14 monotherapy cases were observed amongst DMARD users).

Although the epidemiology of hypertension is well known in the general population(84-86), there is precious little known about essential hypertension in the RA population (with the exception of case series and a few studies of pulmonary hypertension). One study found that the prevalence of pulmonary hypertension amongst RA patients was 21%(87). DMARDs do not appear to confer added risk. In this study, rates of hypertension were lowest amongst leflunomide.

The worldwide incidence of acute pancreatitis appears to be increasing(88), with estimates ranging from one case per million person-years (in women) to over 300 cases per million PY(89-92). This rare condition is seen in rheumatoid arthritis patients(93, 94) and has been associated with gold therapy(95) and mizoribine(96). The rates seen in this study, about 2.5 cases per 1000 person-years, are about 10-times higher than the highest estimates in the literature on the general population. This higher than expected rate is likely due to the fact that the codes used to capture pancreatitis included a code for elevated amylase (a sensitive but not specific marker for pancreatitis).

The epidemiology of interstitial pneumonitis and other interstitial lung diseases is not well characterized. A seminal study was undertaken in New Mexico, and estimated the incidence to be 1.8 cases per 100 000 person-years in men, and 1.4 cases per 100 000 PY in women(97). The total interstitial lung disease (invcluding pulmonary fibrosis and idiopathic pulmonary fibrosis) was 31.5 cases per 100 000 PY in men and 26.2 per 100 000 PY in women.

Methotrexate has been associated with pneumonitis in numerous studies, mostly in case reports but in at least one follow-up study(98). The incidence of pneumonitis was found to be 2.1% in RA patients in a study from Japan(99), 2.1% in an Italian study(100), 3.8% in one French study(101) and 3.2% in a second(102), and 2.8% in an Australian investigation(103). Reviews of methotrexate-induced

pneumonitis note that the prevalence can range from 0.3% to 18%(100) and incidence from 3-5%(104). The incidence in the study cohort was 13 cases per 1000 person-years. The leflunomide rate was 11 per 1000 PY, and the rate amongst DMARD users was 16 per 1000 PY.

Thus, while there should be no expectation that leflunomide should be any more exempt than other DMARDs from the adverse effects seen in this study (and described in the literature), the fact that leflunomide had lower rates of several AEs should be a clear indication that it is quite possibly a safer drug—and at worst, no less safe than comparator agents.

An unexpected observation in this study was that the AE rates associated with monotherapy use were generally higher than those of two-drug therapy. One possible explanation is that the increased incidence may be the result of a 'depletion of susceptibles' effect whereby patients who remain on the drugs are those who can tolerate them while those who are susceptible select themselves out (or are selected out) of the population at risk. Thus, if a certain percent of monotherapy patients were to experience an hepatic event (the susceptibles), say, they would not be available for two-drug therapy and only those who 'survived' the monotherapy would be. These survivors, of course, would be healthier, in the sense that they had not experienced an AE of interest. Furthermore, they may be appropriate candidates for additional therapy. On the other hand, it is possible that the lower rates are due to enhanced efficacy of two- rather than one-drug therapy. There is no evidence from the literature that polypharmacy results in *more* adverse events until the number of drugs used exceeds five.

9. LIMITATIONS/POTENTIAL SOURCES OF BIAS

The data collected and analyzed in this study came from a claims database, and were not designed to be used for research *per se*. Limitations of claims databases include lack of data on over-the-counter medications, potential omission of services provided, potential diagnostic and procedural coding errors, lack of indicators of disease severity, limited clinical detail, little or no data on compliance, potential exposure misclassification, varying and differential lag times for pharmacy and medical claims, and lack of lifetime history of the disease(s) under study along with its treatment (105-108). Furthermore, the fact that the database requires the submission of claims from physicians is dependent upon the willingness of the physician to fill out (electronically or otherwise) such forms with the kind of detail needed in epidemiologic research.

The database itself generally reflects the diversity of the US population in terms of economic, sex, and racial constituencies. It may include information not only on the primary insured person, but on family members of that person as well.

However, the benefit of using claims databases, and the one used in this study in particular, is the recruitment of the largest number of leflunomide patients for study in a short period of time. The assembled cohort is extremely large, with over 40 000 rheumatoid arthritis patients (and 5325 leflunomide users), and this patient population very likely represents thousands of physicians whose beliefs and practices are not amenable to statistical (or any other kind) of adjustment. Thus, while minor adverse events may be undercounted or misattributed (to other drugs or conditions) this is not a main concern as the focus of this study was mostly on serious events.

It is likely that this database may represent a healthier population than the 'general population', as subjects in it are employed and younger in age than the US population. This 'limitation' is countered by the fact of the study design, which relied on internal comparisons within the database.

The validity and reliability of claims data are always a concern for at least two reasons: coding of the data from the original medical records into the computerized database may be faulty, and the source medical record data themselves may be incomplete and inadequate to address the many questions asked. At worst, however, this means that the resulting imprecise data will result in a misconstruction of the drug-induced disease (or its pathophysiology)—but not necessarily bias estimates of risk. Data validity is being examined currently (unfortunately, the results of this exercise are not ready at this time).

This study takes no account of whether a patient's course of therapy was the first the patient received or whether it was the second, third, or fourth trial of a particular drug. Clearly, patients may be at higher (or lower) risk for an AE if there is a history of prior therapy—and, therefore, 'residual effectiveness' (or toxicity) of the previous treatment (109). The absence of such data is not limited to the database chosen for this study but is, rather, a common shortcoming shared by all such databases.

Similarly, the study lacks data on relative RA severity in the different cohorts. There is no way to redress this constraint in the current study.

A recent study found few differences in the 'treatment flow', ie, time on a first line DMARD therapy and time to a second line DMARD treatment (110). For example, current Enbrel users started first line DMARD therapy approximately 45 months after diagnosis and stayed on that therapy for 29 months before switching to a second DMARD; leflunomide users started DMARD therapy 35 months after diagnosis and stayed on that therapy 24 months before switching to a second DMARD, etc. These results confirm other reports (111).

There is little agreement in the literature about the prognosis of RA and, in fact, long-term natural history and results of therapy are strongly influenced by the study designs (112). Even well designed studies that

acknowledge that variations in disease duration and follow up are 'fundamental pitfalls' conclude that disease outcome cannot be predicted accurately with current means (113).

Detection bias, as is true with all forms of selection bias, is an ever-present threat, especially if the outcome of interest is asymptomatic (such as elevations of liver enzymes), the recognition of which is wholly at the discretion of the treating physician, ie, whether he or she decides to do routine liver testing.

Self-selection bias is always a possibility when dealing with claims data from managed care organizations (MCOs) in the United States, which tend to undergo an annual average turnover of about 20%(114). Persons may change health plans several times in a given year, for reasons relating to drug availability (ie, the extent of the formulary), physician availability, and changes in their own health status.

Protopathic bias (115) occurs 'if a particular maneuver was started, stopped, or otherwise changed because of the baseline manifestation caused by a disease or other outcome event', in other words, when the first symptoms of the outcome of interest are the reasons for the use of the treatment that is the subject of a study. This results in a distorted estimate of disease-drug association because the exposure, ie, drug use, started after the occurrence of the event. Due to the simultaneously complicated and incomplete nature of the data used in this study, such a bias is possible.

Several of the outcomes are not as specific as desired, and this will be addressed in future studies. For example, there is no ICD that explicitly refers to elevated liver enzymes, only to 'other nonspecific abnormal serum enzyme levels' and 'nonspecific elevation of levels of transaminase or lactic acid dehydrogenase'. Conversely, the inclusion of 'elevated amylase' as a marker for pancreatitis, is too nonspecific for a detailed interpretation.

It is possible that the assigned exposure windows, based on a time period equivalent to five times the half-life of the specific drugs, resulted in misclassification of adverse events(8). It has been recommended that the exposure window be adjusted for the event of interest, rather than for the elimination period of the drug. This method, however, seems counterintuitive and prone to more misclassification, as the potential effects of various drugs will be forced into a predetermined window. In the current study, the exposure windows for non-leflunomide drugs were kept constant. A sensitivity analysis using 30-, 60-, and 90-day windows for leflunomide was undertaken, a very little difference in adverse event rates was noticed. The 30- and 60-day rates were virtually identical; the 90-day rates were lower, reflecting an inevitable dilution effect. Needless to say, the conclusions drawn from these results remain.

10. CONCLUSIONS

The present investigation, a cohort study encompassing over 83 000 person-years of follow-up, is one of the largest studies of rheumatoid arthritis monotherapy and combination therapy undertaken.

Overall, the incidence rate of any endpoint was statistically significantly lower for leflunomide compared to DMARD and methotrexate use. Similarly, the incidence of any endpoint amongst leflunomide + methotrexate users was lower compared to other two-drug therapy combinations, including DMARD + methotrexate.

We conclude, therefore, that the safety profile of leflunomide is comparable to that of other DMARDs currently used to treat rheumatoid arthritis.

11. REFERENCES

1. **Myers R, Hyslop T, Gerrity M, et al.** Physician intention to recommend complete diagnostic evaluation in clorectal cancer screening. *Cancer Epidemiol Biomarkers Prev*. 1999;8(7):587-93.
2. **Leone F, Grana J, McDermott P, MacPherson S, Hanchak N, Fish J.** Pharmaceutically-based severity stratification of an asthmatic population. *Resp Med*. 1999;159(3):788-93.
3. **Breslow N.** Cohort analysis in epidemiology. In: Atkinson A, Feinberg S, eds. *A celebration of statistics*. New York: Springer-Verlag; 1985.
4. **Breslow N, Day N.** *Statistical methods in cancer research. Vol II: the design and analysis of cohort studies*. Lyon: International Agency for Research on Cancer; 1987.
5. **Frome E.** The analysis of rates using poisson regression models. *Biometrics*. 1983;39(3):665-74.
6. **Holford T.** The analysis of rates and survivorship using log-linear models. *Biometrics*. 1980;36(2):299-305.
7. **Charlson M, Pompei P, Ales K, MacKenzie C.** A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chron Dis*. 1987;40(5):373-83.
8. **van Staa T-P, Abenhaim L, Leufkens H.** A study of the effects of exposure misclassification due to the time-window design in pharmacoepidemiologic studies. *J Clin Epidemiol*. 1994;47(2):183-9.
9. **Wiles N, Symmons D, Harrison B, et al.** Estimating the incidence of rheumatoid arthritis. Trying to hit a moving target? *Arthritis Rheum*. 1999;42(7):1339-46.
10. **Safavi K, Heyse S, Hochberg M.** Estimating the incidence and prevalence of rare rheumatologic diseases: a review of methodology and available data sources. *J Rheumatol*. 1990;17(8):990-3.
11. **Willkens R, Blandau R, Aoyama D, Beasley R.** Studies of rheumatoid arthritis among a tribe of Northwest Indians. *J Rheumatol*. 1976;3(1):9-14.
12. **Harvey J, Arnett F, Bias W, Hsu S, Stevens M.** Heterogeneity of HLA-DR4 in the rheumatoid arthritis of a Chippewa band. *J Rheumatol*. 1981;8(5):797-803.
13. **Del Puente A, Knowler W, Pettitt D, Bennett P.** High incidence and prevalence of rheumatoid arthritis in Pima Indians. *Am J Epidemiol*. 1989;129(6):1170-8.
14. **Hawley D, 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(7):1008-14.
15. **Pincus T.** Rheumatoid arthritis: disappointing long-term outcomes despite successful short-term clinical trials. *J Clin Epidemiol*. 1988;41(11):1037-41.
16. **Boelsterli U, Zimmerman H, Kretz-Rommel A.** Idiosyncratic liver toxicity of nonsteroidal antiinflammatory drugs: molecular mechanisms and pathology. *Crit Rev Toxicol*. 1995;25(3):207-35.
17. **Dahl S, Ward J.** Pharmacology, clinical efficacy, and adverse effects of the nonsteroidal anti-inflammatory agent beonxaprofen. *Pharmacotherapy*. 1982;2(6):354-66.
18. **Rabinovitz M, Van Thiel D.** Hepatotoxicity of nonsteroidal anti-inflammatory drugs. *Am J Gastroenterol*. 1992;87(12):1696-1704.
19. **Brass E.** Hepatic toxicity of antirheumatic drugs. *Cleve Clin J Med*. 1993;60(6):466-72.
20. **Friis H, Andreasen P.** Drug-induced hepatic injury: an analysis of 1100 cases reported to the Danish Committee on Adverse Drug Reactions between 1978 and 1987. *J Intern Med*. 1992;232(2):133-8.
21. **Carson J, Strom B, Duff A, Gupta A, Das K.** Safety of nonsteroidal anti-inflammatory drugs with respect to acute liver disease. *Arch Intern Med*. 1993;153:1331-6.
22. **Garcia Rodriguez L, Perez Gutthann S, Walker A, Lueck L.** The role of non-steroidal anti-inflammatory drugs in acute liver injury. *BMJ*. 1992;305:865-8.
23. **Jick H, Derby L, Garcia Rodriguez L, Jick S, Dean A.** Liver disease associated with diclofenac, naproxen, and piroxicam. *Pharmacotherapy*. 1992;12(3):207-12.
24. **Garcia Rodriguez L, Williams R, Derby L, Dean A, Jick H.** Acute liver injury associated with nonsteroidal anti-inflammatory drugs and the role of risk factors. *Arch Intern Med*. 1994;154:311-6.
25. **Walker A.** Quantitative studies of the risk of serious hepatic injury in persons using nonsteroidal antiinflammatory drugs. *Arthritis Rheum*. 1997;40(2):201-8.
26. **Walker A, Bortnichak E, Lanza L, Yood R.** The infrequency of liver function testing in patients using nonsteroidal anti-inflammatory drugs. *Arch Fam Med*. 1995;4:24-9.

27. **Manoukian A, Carson J.** Nonsteroidal anti-inflammatory drug-induced hepatic disorders: incidence and prevalence. *Drug Safety*. 1996;15(1):64-71.
28. **Fry S, Seeff L.** Hepatotoxicity of analgesics and anti-inflammatory agents. *Gastroenterol Clin North Am*. 1995;24(4):875-905.
29. **Figueras A, Estevez F, Laporte J-R.** New drugs, new adverse drug reactions, and bibliographic databases. *Lancet*. 1999:1447-8.
30. **McCormick P, Kennedy F, Curry M, Traynor O.** Cox 2 inhibitor and fulminant hepatic failure. *Lancet*. 1999;353:40-1.
31. **Galan M, Gordon S, Silverman A.** Celecoxib-induced cholestatic hepatitis. *Ann Intern Med*. 2001;134(3):254.
32. **Nachimuthu S, Volfinzon L, Gopal L.** Acute hepatocellular and cholestatic injury induced in a patient taking celecoxib. *Postgrad Med J*. 2000;77:548-50.
33. **Hawkey C, Lanas A.** Doubt and certainty about nonsteroidal anti-inflammatory drugs in the year 2000: a multidisciplinary expert statement. *Am J Med*. 2001;110 (suppl 1A):79S-100S.
34. **Emery P.** Cyclooxygenase-2: a major therapeutic advance? *Am J Med*. 2001;110 (suppl 1A):42S-5S.
35. **Schnitzer T.** Cyclooxygenase-s-specific inhibitors: are they safe? *Am J Med*. 2001;110 (suppl 1A):46S-49S.
36. **Brater D, Harris C, Redfern J, Gertz B.** Renal effects of cox-2-selective inhibitors. *Nephrology*. 2001;21:1-15.
37. **Perazella M, Tray K.** Selective cyclooxygenase-2 inhibitors: a pattern of nephrotoxicity similar to traditional nonsteroidal anti-inflammatory drugs. *Am J Med*. 2001;111:64-7.
38. **Eras J, Perazella M.** NSAIDs and the kidney revisited: are selective cyclooxygenase-2 inhibitors safe? *Am J Med Sci*. 2001;321(3):181-90.
39. **Halter F, Tarnawski A, Schmassmann A, Peskar B.** Cyclooygenase 2 -- implications on maintenance of gastric mucosal integrity and ulcer healing: controversial issues and perspectives. *Gut*. 2001;49:443-53.
40. **Cleland L, James M, Stamp L, Penglis P.** Cox-2 inhibition and thrombotic tendency: a need for surveillance. *Med J Aust*. 2001;175:214-7.
41. **Raisz L.** Potential impact of selective cyclooxygenase-2 inhibitors on bone metabolism in health and disease. *Am J Med*. 2001;110 (supp 3A):43S-45S.
42. **n.a.** Celecoxib in osteoarthritis and rheumatoid arthritis: new preparation. As disappointing as rofecoxib. *Prescrire Int*. 2001;10(52):46-9.
43. **Mol M, Ng W, Yuen M, Wong R, Lau C.** Safety of disease modifying anti-rheumatic agents in rheumatoid arthritis patients with chronic viral hepatitis. *Clin Exp Rheumatol*. 2000;18:363-8.
44. **Grove M, Hassell A, Hay E, Shdforth M.** Adverse reactions to disease-modifying anti-rheumatic drugs in clinical practice. *Q J Med*. 2001;94:309-19.
45. **Galindo-Rodriguez G, Avina-Zubieta J, Russell A, Suarez-Almazour M.** Disappointing longterm results with disease modifying antirheumatic drugs. A practice based study. *J Rheumatol*. 1999;26(11):2337-43.
46. **Beyeler C, Reichen J, Thomann S, Lauterburg B, Gerber N.** Quantitative liver function in patients with rheumatoid arthritis treated with low-dose methotrexate: a longitudinal study. *Br J Rheumatol*. 1997;36:338-44.
47. **Singh G, Fries J, Williams C, Zatarain E, Spitz P, Bloch D.** Toxicity profiles of disease modifying antirheumatic drugs in rheumatoid arthritis. *J Rheumatol*. 1991;18(2):188-94.
48. **Bailey A.** Biochemistry of well populations. *Lancet*. 1974;2(7894):1436-9.
49. **Bates B, Yellin J.** The yield of multiphasic screening. *JAMA*. 1972;222(1):74-8.
50. **Lawrence C, Trewin V.** The construction of biochemical reference ranges and the identification of possible adverse drug reactions in the elderly. *Statist Med*. 1991;10:831-7.
51. **Kools A, Bloomer J.** Abnormal liver function tests. How to assess their importance in asymptomatic patients. *Postgrad Med*. 1987;81(6):45-51.
52. **McKenna J, Moskovitz M, Cox J.** Abnormal liver function tests in asymptomatic patients. . 1989;39:117-26.
53. **Younossi Z.** Evaluating asymptomatic patients with mildly elevated liver enzymes. *Clev Clin J Med*. 1998;65(3):150-8.
54. **Duh M-S, Walker A, Kronlund K.** Descriptive epidemiology of acute liver enzyme abnormalities in the general population of central Massachusetts. *Pharmacoepidemiol Drug Safety*. 1999;8:275-83.

55. **Walker A, Cavanaugh R.** The occurence of new hepatic disorders in a defined population. *Post Marketing Surveillance*. 1992;6:107-17.
56. **Wolfe F.** The natural history of rheumatoid arthritis. *J Rheumatol*. 1996;23 (suppl 44):13-22.
57. **Pincus T, Callahan L.** What is the natural history of rheumatoid arthritis? *Rheum Clin North Am*. 1993;19(1):123-51.
58. **Movitt E, Davis A.** Liver biopsy in rheumatoid arthritis. *Am J Med*. 1953;226:516-20.
59. **Lefkovits A, Farrow I.** The liver in rheumatoid arthritis. *Ann Rheum Dis*. 1955;14:162-8.
60. **Ruderman E, Crawford J, Maier A, Liu J, Gravallese E, Weinblatt M.** Histologic liver abnormalities in an autopsy series of patients with rheumatoid arthritis. *Br J Rheumatol*. 1997;36(2):210-3.
61. **Mills P, MacSween R, Dick W, More I, Watkinson G.** LIver disease in rheumatoid arthritis. *Scott Med J*. 1980;25(1):18-22.
62. **Mills P, Sturrock R.** Clinical associations between arthritis and liver disease. *Ann Rheum Dis*. 1982;41(3):295-307.
63. **Saway P, Prasthofer E, Barton J.** Prevalence of granular lymphocyte proliferation in patients with rheumatoid arthritis and neutropenia. *Am J Med*. 1989;86(3):303-7.
64. **Yong A, Goh A, Rahman M, Menon J, Purushothaman V.** Epidemiology of aplastic anaemia in the state of Sabah, Malaysia. *Med J Malayasia*. 1998;53(1):59-62.
65. **Shapiro S, Issaragrisil S, Kaufman D, et al.** Agranulocytosis in Bangkok, Thailand: a predominantly drug-induced disease with an unusually low incidence. Aplastic Anemia Study Group. *Am J Trop Med Hyg*. 1999;60(4):573-7.
66. **Issaragrisil S, Leaverton P, Chansung K, et al.** Regional patterns in the incidence of aplastic anemia in Thailand. The Aplastic Anemia Study Group. *Am J Hematol*. 1999;61(3):164-8.
67. **Larregina A, Aggio M, Alvarez R.** Incidence and probable etiology of toxic agranulocytosis in a definite population in the province of Buenos Aires (1963-1976). *Medicina (B Aires)*. 1994;54(1):13-6.
68. **Strom B, Carson J, Schinnar R, Snyder E, Shaw M.** Descriptive epidemiology of agranulocytosis. *Arch Intern Med*. 1992;152(7):1475-80.
69. **Rawson N, Harding S, Malcolm E, Lueck L.** Hospitalizations for aplastic anemia and agranulocytosis in Saskatchewan: incidence and associations with antecedent prescription drug use. *J Clin Epidemiol*. 1998;51(12):1343-55.
70. **Baumelou E, Guiguet M, Mary J.** Epidemiology of aplastic anemia in France: a case-control study. I. Medical history and medication use. The French Cooperative Group for Epidemiological Study of Aplastic Anemia. *Blood*. 1993;81(6):1471-8.
71. **Gibson J, McGirr E, York J, Kronenberg H.** Aplastic anemia in association with gold therapy for rheumatoid arthritis. *Aust N Z J Med*. 1983;13(2):130-4.
72. **Canvin J, el-Gabalawy H, Chalmers I.** Fatal agranulocytosis with sulfasalazine therapy in rheumatoid arthritis. *J Rheumatol*. 1993;20(5):909-10.
73. **Pertuiset E, Liote F, Chevret S, Bardin T, Dryll A, Kuntz D.** Therapeutic maintenance and tolerance of sulfasalazine in rheumatoid polyarthritis. Retrospective study of 95 patients. *Rev Rhum Mal Osteoartic*. 1992;59(9):545-52.
74. **McKendry R, Dale P.** Adverse effects of low dose methotrexate therapy in rheumatoid arthritis. *J Rheumatol*. 1993;20(11):1850-6.
75. **Scott D, Dacre J.** Adverse reactions to sulfasalazine: the British experience. *J Rheumatol Suppl*. 1988;16:17-21.
76. **Schattenkirchner M, Missler B, Mulz D.** Postmarketing experience with auranofin in the Federal Republic of Germany. *Scand J Rheumatol Suppl*. 1986;63:57-66.
77. **Hirakawa S, Ohtsu T, Ojima Y, et al.** Acute myocardial infarction following toxic epidermal necrolysis? *Clin Exp Dermatol*. 1999;24(5):375-8.
78. **Roupe G, Ahlmen M, Fagerberg B, Suurkula M.** Toxic epidermal necrolysis with extensive mucosal erosions of the gastrointestinal and respiratory tracts. *Int Arch Allergy Appl Immunol*. 1986;80(2):145-51.
79. **O'Sullivan M, Hanly J, Molloy M.** A case of toxic epidermal necrolysis secondary to indomethacin. *Br J Rheumatol*. 1983;22(2):47-9.
80. **Laurencin C, Horan R, Senatus P, Wheeler C, Lipson S.** Stevens-Johnson-type reaction with vancomycin treatment. *Ann Pharmacother*. 1992;26(12):1520-1.

81. **Rzany B, Mockenhaupt M, Baur S, et al.** Epidemiology of erythma exsudativum multiforme majus, Stevens-Johnson syndrome and toxic epidermal necrolysis in Germany (1990-1992): structure and results of a population-based registry. *J Clin Epidemiol*. 1996;49(7):769-73.
82. **Sane S, Bhatt A.** Stevens-Johnson syndrome and toxic epidermal necrolysis--challenges of recognition and management. *J Assoc Physicians India*. 2000;48:999-1003.
83. **Wolkenstein P, Revuz J.** Drug-induced severe skin reactions. Incidence, management and prevention. *Drug Safety*. 1995;13(1):56-68.
84. **Kornitzer M, Dramaix M, De Backer G.** Epidemiology of risk factors for hypertension: implications for prevention and therapy. *Drugs*. 1999;57(5):695-712.
85. **Burt V, Whelton P, Roccella E, et al.** Prevalence of hypertension in the US adult population: results from the Third National Health and Nutrition Examination Survey, 1988-1991. *Hypertens*. 1995;25(3):305-13.
86. **Burt V, Cutler J, Higgins M, et al.** Trends in the prevalence, awareness, treatment and control of hypertension in the US adult population: data from the health examination surveys, 1960 to 1991. *Hypertens*. 1995;26(1):60-9.
87. **Dawson J, Goodson N, Graham D, Lynch M.** Raised pulmonary artery pressures measured with Doppler echocardiography in rheumatoid arthritis patients. *Rheumatology (Oxford)*. 2000;39(12):1320-5.
88. **Corfield A, Cooper M, Williamson R.** Acute pancreatitis: a lethal disease of increasing incidence. *Gut*. 1985;26:724-9.
89. **Dubagunta S, Still C, Komar M.** Acute pancreatitis. *J Am Osteopath Assoc*. 2001;101(4 (suppl 1)):S6-9.
90. **Eland I, Sturkenboom M, Wilson J, Stircker B.** Incidence and mortality of acute pancreaitis between 1985 and 1995. *Scand J Gastroenterol*. 2000;35(10):1110-6.
91. **Giggs J, Bourke J, Katschinski B.** The epidemiology of primary acute pancreatitis in Greater Nottingham: 1969-1983. *Soc Sci Med*. 1988;26(1):79-89.
92. **Thomson H.** Acute pancreatitis in North and North-East Scotland. *J Royal Coll Surg*. 1985;30(2):104-11.
93. **Bely M, Apathy A, Pinter T, Ratko I.** Pancreatitis in rheumatoid arthritis found in autopsy material. *Morphol Igazsagugyi Orv Sz*. 1989;29(3):213-21.
94. **Oishi K, Wada J, Nagake Y, et al.** Fatal pancreatitis associated with systemic amyloidosis in a rheumatoid arthritis patient. *J Med*. 2000;31(5-6):303-10.
95. **Ben-Ami H, Pollack S, Nagachandran P, Lashevsky I, Yarnitsky D, Edoute Y.** Reversible pancreatitis, hepatitis, and peripheral polyneuropathy associated with parenteral gold therapy. *J Rheumatol*. 1999;26(9):2049-50.
96. **Yutsudo Y, Hasegawa Y, Takada T, et al.** A case of rheumatoid arthritis (RA) with drug-induced acute pancreatitis due to mizoribine. *Ryumachi*. 1997;37(4):564-7.
97. **Coultas D, Zumwalt R, Black W, Sobonya R.** The epidemiology of interstitial lung diseases. *Am J Respir Crit Care Med*. 1994;150(4):967-72.
98. **Schnabel A, Herlyn K, Burchardi C, Reinhold-Keller E, Gross W.** Long-term tolerability of methotrexate at doses exceeding 15 mg per week in rheumatoid arthritis. *Rheumatol Int*. 1996;15(5):195-200.
99. **Ohosone Y, Okano Y, Kameda H, et al.** Clinical characteristics of patients with rheumatoid arthritis and methotrexate induced pneumonitis. *J Rheumatol*. 1997;24(12):2299-303.
100. **Salaffi F, Manganelli P, Carotti M, Subiaco S, Lamanna G, Cervini C.** Methotrexate-induced pneumonitis in patients with rheumatoid arthritis and psoriatic arthritis: report of five cases and review of the literature. *Clin Rheum*. 1997;16(3):296-304.
101. **Hilliquin P, Renoux M, Perrot S, Puechal X, Menkes C.** Occurrence of pulmonary complications during methotrexate therapy in rheumatoid arthritis. *Br J Rheumatol*. 1996;35(5):441-5.
102. **Cottin V, Tebib J, Massonnet B, Souquet P, Bernard J.** Pulmonary function in patients receiving long-term low-dose methotrexate. *Chest*. 1996;109(4):933-8.
103. **Carroll G, Thomas R, Phatouros C, et al.** Incidence, prevalence and possible risk factors for pneumonitis in patients with rheumatoid arthritis receiving methotrexate. *J Rheumatol*. 1994;21(1):51-4.
104. **Sharma A, Provenzale D, McKusick A, Kaplan M.** Interstitial pneumonitis after low-dose methotrexate therapy in primary biliary cirrhosis. *Gastroenterology*. 1994;107(1):266-70.

105. **Armstrong E, Manuchehri F.** Ambulatory care databases for managed care organizations. *Am J Health Syst Pharm*. 1997;54(17):1973-83.
106. **Deyo R, Taylor V, Diehr P, et al.** Analysis of automated administrative and survey databases to study patterns and outcomes of care. *Spine*. 1994;19 (suppl 18):2083S-2091S.
107. **Motheral B, Fairman K.** The use of claims databases for outcomes research: rationale, challenges, and strategies. *Clin Ther*. 1997;19(2):346-66.
108. **Shatin D, Drinakrd C, Stergachis A.** UnitedHealth Group. In: Strom B, ed. *Pharmacoepidemiology*. 3rd ed. Chichester: John wiley & Sons; 2000:295-306.
109. **Fries J, Williams C, Singh G, Ramey D.** Response to therapy in rheumatoid arthritis is influenced by immediately prior therapy. *J Rheumatol*. 1997;24(5):838-44.
110. **TNSHealthcare.** Arava performance tracking and rheumatoid arthritis DMARD patient segmentation: international findings. 2001.
111. **Suarez-Almazor M, Soskolne C, Saunders L, Russell A.** Outcome in rheumatoid arthrits. A 1985 inception cohort study. *J Rheumatol*. 1994;21(8):1438-46.
112. **Pincus T.** Assessment of long-term outcomes of rheumatoid arthritis. How choices of measures and study designs may lead to apparently different conclusions. *Rheum Dis Clin North Am*. 1995;21(3):619-54.
113. **Mottonen T, Paimela L, Leirisalo-Repo M, Kautiainen H, Ilonen J, Hannonen P.** Only high disease activity and positive rheumatoid factor indicate poor prognosis in patients with early rheumatoid arthritis treated with 'sawtooth' strategy. *Ann Rhem Dis*. 1998;57:533-9.
114. **Keister L.** Switching doctors, switching plans: the 'revolving door' problem. *Managed Care*. 1996;5(5):n.p.
115. **Feinstein A.** *Clinical epidemiology: the architecture of clinical research*. Philadelphia: WB Saunders; 1985.
116. **Lawrence R, Hochberg M, Kelsey J, et al.** Estimates of the prevalence of selected arthritic and musculoskeletal diseases in the United States. *J Rheumatol*. 1989;16:427-41.
117. **Mikkelsen W, Dodge H, Duff I.** Estimates of the prevalence of rheumatic diseases in the population of Tecumseh, Michigan, 1959-1960. *J Chronic Dis*. 1967;20:351-69.
118. **Cathcart E, O'Sullivan J.** Rheumatoid arthritis in a New England town: a prevalence study in Sudbury, Massachusetts. *N Engl J Med*. 1970;282:421-4.
119. **Cunningham L, Kelsey J.** Epidemiology of musculoskeletal impairments and associated disability. *Am J Public Health*. 1984;74:574-9.
120. **Linos A, Worthington J, O'Fallon W, Kurland L.** The epidemiology of rheumatoid arthritis in Rochester, Minnesota: a study of incidence, prevalence, and mortality. *Am J Epidemiol*. 1980;111:87-98.
121. **Gabriel S, Crowson C, O'Fallon W.** The epidemiology of rheumatoid arthritis in Rochester, Minnesota, 1955-1985. *Arthritis Rheum*. 1999;42(3):415-20.
122. **Stojanovic R, Vlajinac H, Palic-Obradovic D, Janosevic S, Adanja B.** Prevalence of rheumatoid arthritis in Belgrade, Yugoslavia. *Br J Rheumatol*. 1998;37(7):729-32.
123. **Shichikawa K, Inoue K, Hirota S, et al.** Changes in the incidence and prevalence of rheumatoid arthritis in Kamitonda, Wakayama, Japan, 1965-1996. *Ann Rheum Dis*. 1999;58:751-6.
124. **Carmona L, Ballina J, Gabriel R, Laffon A.** The burden of musculoskeletal diseases in the general population of Spain: results from a national survey. *Ann Rheum Dis*. 2001;60:1040-5.
125. **Farooqi A, Gibson T.** Prevalence of the major rheumatic disorders in the adult population of North Pakistan. *Br J Rheumatol*. 1998;37(5):491-5.
126. **Riise T, Jacobsen B, Gran J.** Incidence and prevalence of rheumatoid arthritis in the county of Troms, northern Norway. *J Rheumatol*. 2000;27(6):1386-9.
127. **Drosos A, Alamanos I, Voulgari P, et al.** Epidemiology of adult rheumatoid arthritis in northwest Greece 1987-1995. *J Rheumatol*. 1997;24:2129-33.
128. **Simonsson M, Bergman S, Jacobsson L, Petersson I, Svensson B.** The prevalence of rheumatoid arthritis in Sweden. *Scand J Rheumatol*. 1999;28(6):340-3.
129. **Spector T, Hart D, Powell R.** Prevalence of rheumatoid arthritis and rheumatoid factor in women: evidence for a secular decline. *Ann Rheum Dis*. 1993;52(4):254-7.
130. **Cimmino M, Parisi M, Moggiana G, Mela G, Accardo S.** Prevalence of rheumatoid arthritis in Italy: the Chiavari study. *Ann Rheum Dis*. 1998;57:315-8.
131. **Saraux A, Guedes C, Allain J, et al.** Prevalence of rheumatoid arthritis and spondylarthropathy in Brittany, France. *J Rheumatol*. 1999;26(12):2622-7.

132. **Dugowson C, Koepsell T, Voight L, Bley L, Nelson J, Daling J.** Rheumatoid arthritis in women. Incidence rates in group health cooperative, Seattle, Washington, 1987-1989. *Arthritis Rheum*. 1991;34(12):1502-7.
133. **Chan K, Felson D, Yood R, Walker A.** Incidence of rheumatoid arthritis in central Massachusetts. *Arthritis Rheum*. 1993;36(12):1691-6.
134. **Guillemin F, Briancon S, Klein J, Sauleau E, Pourel J.** Low incidence of rheumatoid arthritis in France. *Scand J Rheumatol*. 1994;23(5):264-8.
135. **Kaipiainen-Seppanen O, Aho K, Isomaki H, Laakso M.** Incidence of rheumatoid arthritis in Finland during 1980-1990. *Ann Rheum Dis*. 1996;55(9):608-11.
136. **Uhlig T, Kvien T, Glennas A, Smedstad L, O F.** The incidence and severity of rheumatoid arthritis, results from a county register in Oslo, Norway. *J Rheumatol*. 1998;25(6):1078-84.
137. **Symmons D, Barrett E, Bankhead C, Scott D, Silman A.** The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. *Br J Rheumatol*. 1994;33(8):735-9.

12. TABLES

TABLE 1 SAMPLE CODES FOR HEPATIC EVENTS

Diagnosis	ICD-9CM
Acute or Subacute Liver Necrosis*	570
Hepatitis, Noninfectious toxic*	573.3
Jaundice	782.4
Cirrhosis of liver, no alcohol	571.5
Biliary cirrhosis	571.6
Hepatitis, noninfectious	573.3
Other specified liver disorder	573.8
Unspecified liver disorder	573.9
Hepatic coma	572.2
Elevated transaminase/LAH	790.4

* cases of particular interest

TABLE 2 DEMOGRAPHIC CHARACTERISTICS OF THE PATIENTS

AGE GROUP	MALE (%)	FEMALE (%)	TOTAL (%)
18-30	408	1554	1962 (4.83)
31-50	3341	9951	13 292 (32.74)
51-64	3499	9485	12 984 (31.99)
65+	3598	8758	12 356 (30.44)
TOTAL	10 846 (26.72)	29 748 (73.28)	40 594

TABLE 3 DESCRIPTIVE STATISTICS OF PERSON-TIME EXPOSURES ACROSS SELECTED DRUG GROUPS

EXPOSURE GROUPS	PERSON-YEAR EXPOSURE	MEAN EXPOSURE TIME (DAYS)	PATIENTS ON THERAPY
LEF	4214	584.6	2633
MTX	10682	410.0	9514
DMARD	31158	765.8	14861
non-DMARD	11259	377.4	10896
LEF + MTX	1415	214.6	2408
LEF + DMARD	5551	753.2	2692
DMARD + MTX	18864	789.7	8725
TOTAL	83 543		

TABLE 4 INCIDENCE RATES (PER 1000 P-Y) OF ANY ENDPOINT

N = NUMBER OF OBSERVED EVENTS

EXPOSURE GROUP (N)	CRUDE RATE (95% CI)	ADJUSTED RATE (95% CI)
LEF MONO (465)	110.35 (100.76, 120.85)	94.06 (84.41, 104.81)
MTX MONO (1789)	167.48 (159.89, 175.42)	145.01 (136.29, 154.30)
DMARD MONO (5475)	175.72 (171.12, 180.44)	143.68 (137.40, 150.25)
LEF + MTX (72)	50.90 (40.40, 64.12)	42.82 (32.81, 55.88)
LEF + DMARD (370)	66.66 (60.20, 73.81)	58.69 (52.02, 66.20)
MTX + DMARD (1512)	80.15 (76.21, 84.29)	69.50 (65.03, 74.28)
NO DMARD (4934)	438.25 (426.19, 450.65)	382.33 (365.80, 399.61)

TABLE 5 INCIDENCE RATES (PER 1000 P-Y) OF HEPATIC EVENTS (HEPATIC NECROSIS, BILLIARY CIRRHOSIS, LIVER CIRRHOSIS, HEPATITIS, OTHER SPECIFIED LIVER DISORDER, UNSPECIFIED LIVER DISORDER, AND ELEVATION OF ENZYMES); N = NUMBER OF OBSERVED EVENTS

EXPOSURE GROUP (N)	CRUDE RATE (95% CI)	ADJUSTED RATE (95% CI)
LEF MONO (22)	5.22 (3.04, 7.40)	4.09 (2.38, 7.02)
MTX MONO (87)	8.14 (6.43, 9.86)	6.89 (5.10, 9.30)
DMARD MONO (246)	7.89 (6.91, 8.88)	4.21 (3.31, 5.33)
LEF + MTX (7)	4.95 (1.28, 8.61)	4.57 (1.88, 11.11)
LEF + DMARD (19)	3.42 (1.88, 4.96)	2.61 (1.45, 4.68)
MTX + DMARD (64)	3.39 (2.56, 4.22)	2.92 (2.09, 4.08)
NO DMARD (199)	17.32 (14.89, 19.75)	13.02 (10.37, 16.33)

TABLE 6 RATES OF INDIVIDUAL LIVER EVENTS

RATES (WITH NUMBER OF EVENTS) PRESENTED PER 10 000 PERSON-YEARS, by EVENT (ICD-9 CODE)
GREYED OUT CELLS = NO EVENTS

	NECROSIS (570)	HEPATIC COMA (572.2)	BILIARY CIRRHOSIS (571.6)	CIRRHOSIS (571.5)	JAUNDICE (782.4)
LEF	2.37 (1)				2.37 (1)
MTX	0.64 (2)			0.96 (3)	1.28 (4)
DMARD	6.55 (7)	4.68 (5)	8.43 (9)	23.40 (25)	6.55 (7)
LEF + MTX					
LEF + DMARD		1.80 (1)			1.80 (1)
MTX + DMARD			0.53 (1)	3.18 (6)	1.59 (3)
NO DMARD	1.78 (2)	1.78 (2)	3.55 (4)	7.11 (8)	7.11 (8)

	NON-INFECTIOUS HEPATITIS (573.3)	CHRONIC LIVER (571.9)	UNSPECIFIED (573.9)	ELEVATED ENZYMES (790.4)
LEF	21.36 (9)		7.11 (3)	14.24 (6)
MTX	9.95 (31)		8.34 (26)	3.53 (11)
DMARD	71.15 (76)	2.81 (3)	30.89 (33)	48.68 (52)
LEF + MTX	35.35 (5)			14.14 (2)
LEF + DMARD	10.81 (6)		10.81 (6)	5.40 (3)
MTX + DMARD	6.89 (13)	0.53 (1)	10.07 (19)	7.95 (15)
NO DMARD	52.40 (59)	3.55 (4)	37.31 (42)	39.08 (44)

TABLE 7 INCIDENCE RATES (PER 1000 P-Y) OF HEMATOLOGIC EVENTS (APLASTIC ANEMIA, PANCYTOPENIA); N = NUMBER OF OBSERVED EVENTS

EXPOSURE GROUP (N)	RATE (95% CI)	ADJUSTED RATE (95% CI)
LEF MONO (3)	0.71 (0, 1.52)	NC*
MTX MONO (8)	0.75 (0.23, 1.27)	NC
DMARD MONO (42)	1.35 (0.94, 1.76)	NC
LEF + MTX (0)	0	NC
LEF + DMARD (1)	0.18 (0, 0.53)	NC
MTX + DMARD (13)	0.69 (0.31, 1.06)	NC
NO DMARD (19)	1.69 (0.93, 2.45)	NC

* NC = not calculable

TABLE 8 INCIDENCE RATES (PER 100 P-Y) OF SEVERE SKIN REACTIONS (ERYTHEMA MULTIFORMAE, STEVENS-JOHNSON SYNDROME, TOXIC EPIDERMAL NECROLYSIS)

EXPOSURE GROUP	RATE (95% CI)	ADJUSTED RATE (95% CI)
LEF MONO (0)	0	NC*
MTX MONO (6)	0.56 (0.11, 1.01)	NC
DMARD MONO (16)	0.51 (0.26, 0.77)	NC
LEF + MTX (0)	0	NC
LEF + DMARD (1)	0.18 (0, 0.53)	NC
MTX + DMARD (3)	0.16 (0, 0.34)	NC
NO DMARD (1)	0.09 (0, 0.26)	NC

* NC = not calculable

TABLE 9 INCIDENCE RATES (PER 100 P-Y) HYPERTENSION (ESSENTIAL HYPERTENSION, EXCLUDING PULMONARY)

EXPOSURE GROUP (N)	RATE (95% CI)	ADJUSTED RATE (95% CI)
LEF MONO (222)	52.68 (45.75, 59.61)	33.17 (27.99. 39.31)
MTX MONO (868)	81.26 (75.85, 86.66)	51.15 (45.66, 57.29)
DMARD MONO (2354)	75.55 (72.50, 78.60)	47.57 (43.18, 52.42)
LEF + MTX (32)	22.61 (14.78, 30.45)	13.05 (8.67, 19.65)
LEF + DMARD (167)	30.08 (25.52, 34.65)	19.60 (16.20, 23.71)
MTX + DMARD (682)	36.15 (33.44, 38.87)	22.76 (20.16, 25.68)
NO DMARD (2862)	254.20 (244.88, 263.51)	157.53 (143.35, 173.09)

TABLE 10 INCIDENCE RATES (PER 100 P-Y) PNEUMONITIS

EXPOSURE GROUP (N)	RATE (95% CI)	ADJUSTED RATE (95% CI)
LEF MONO (56)	13.29 (10.23, 17.27)	11.03 (7.93, 15.34)
MTX MONO (182)	17.04 (14.73, 19.70)	13.09 (10.62, 16.14)
DMARD MONO (719)	23.08 (21.45, 24.83)	15.91 (13.77, 18.37)
LEF + MTX (12)	8.48 (4.82, 14.94)	6.67 (3.31, 13.45)
LEF + DMARD (62)	11.17 (8.71, 14.33)	9.44 (6.88, 12.97)
MTX + DMARD (184)	9.75 (8.44, 11.27)	7.60 (6.16, 9.37)
NO DMARD (321)	28.51 (25.56, 31.81)	21.54 (18.16, 25.55)

TABLE 11 INCIDENCE RATES (PER 1000 P-Y) OF PANCREATITIS

EXPOSURE GROUP (N)	RATE (95% CI)	ADJUSTED RATE (95% CI)
LEF MONO (8)	1.90 (0.58, 3.21)	1.22 (0.45, 3.34)
MTX MONO (21)	1.97 (1.13, 2.81)	2.07 (1.21, 3.55)
DMARD MONO (99)	3.18 (2.55, 3.80)	2.40 (1.68, 3.43)
LEF + MTX (1)	0.71 (0, 2.09)	NC*
LEF + DMARD (8)	1.44 (0.44, 2.44)	1.63 (0.75, 3.57)
MTX + DMARD (23)	1.22 (0.72, 1.72)	1.24 (0.73, 2.11)
NO DMARD (53)	4.71 (3.44, 5.97)	3.64 (2.40, 5.53)

* NC = not calculable

TABLE 12 INCIDENCE RATES (PER 100 P-Y) OF RESPIRATORY EVENTS (ACUTE BRONCHITIS, INFLUENZA, ACUTE LARYNGOPHARYNGITIS, BRONCHITIS)

EXPOSURE GROUP (N)	RATE (95% CI)	ADJUSTED RATE (95% CI)
LEF MONO (94)	22.31 (17.80, 26.82)	20.12 (15.98, 25.32)
MTX MONO (436)	40.82 (36.99, 44.65)	38.93 (34.60, 43.80)
DMARD MONO (1235)	39.64 (36.99, 44.65)	36.91 (34.01, 40.06)
LEF + MTX (17)	12.01 (6.30, 17.73)	11.82 (7.10, 19.66)
LEF + DMARD (68)	12.25 (.34, 15.16)	11.62 (8.91, 15.14)
MTX + DMARD (365)	19.35 (17.36, 21.33)	18.95 (16.73, 21.47)
NO DMARD (1003)	89.08 (83.57, 94.60)	88.99 (81.85, 96.75)

TABLE 13 PREVALENCE OF RA ACROSS DIFFERENT STUDIES

STUDY SITE, YEAR(S)	PREVALENCE PER 1 000	DIAGNOSIS INCLUDED:	REFERENCE
US, 1960-62	7 (men) 16 (women)	Rheumatoid factor serology, radiographs	(116)
Tecumseh, 1959-60	3 (men) 7 (women)	Rheumatoid factor serology	(117)
Sudbury, 1960-62	5 (men) 15 (women)	Exam, rheumatoid factor serology, radiography	(118)
US, 1971-75	5 (men) 10 (women)	Exam (clinical diagnosis)	(119)
Rochester, 1950-74	6.6 (men) 13.1 (women)	ACR criteria (1958)	(120)
Rochester, 1955-85	7.4 (men) 13.7 (women)	ACR criteria (1987)	(121)
Belgrade, 1990-91	0.9 (men) 2.9 (women)	ARA criteria (1987), exam, rheumatoid factor, x-ray	(122)
Kamitonda, 1996	1.1 (men) 2.4 (women)	ARA criteria (1961 Rome), exam, serology, x-ray	(123)
Spain, n.d.	5 (sexes combined)	ACR criteria (1992)	(124)
North Pakistan, 1997	1-5 (sexes combined)	ARA criteria (1987)	(125)
Norway, 1987-96	3.0 (men) 6.3 (women)	ARA criteria (1987)	(126)
Greece, 1995	2.1 (men) 4.8 (women)	ARA criteria (1987)	(127)
Sweden, n.d.	5.1 (sexes combined)	ARA criteria (1987)	(128)
UK, n.d.	12 (women)	ARA criteria (1958)	(129)
Italy, 1991-2	1.3 (men) 5.1 (women)	ARA criteria (1987)	(130)
Brittany, n.d.	3.2 (men) 6.2 (women)	Interview with patient, physician	(131)

TABLE 14 INCIDENCE RATES OF RA ACROSS DIFFERENT STUDIES

STUDY SITE, YEAR(S)	INCIDENCE, PER 100 000 PER YEAR	DIAGNOSIS INCLUDED:	REFERENCE
England, 1990	30.8 (women) 12.7 (men)	ACR criteria (1987)	(9)
Seattle, 1987-89	27.9 (women)	ACR criteria (1987)	(132)
Rochester, 1950-74	65.7 (women) 28.1 (men)	ACR criteria (1958)	(120)
Rochester, 1955-85	98.1 (women) 49.7 (men)	ACR criteria (1987)	(121)
Massachusetts, 1987-90	60 (women) 22 (men)	ACR criteria (1958)	(133)
France, 1986-89	12.7 (women) 4.7 (men)	ACR criteria (1958)	(134)
Finland, 1985	39 (sexes combined)	ACR criteria (1987)	(135)
Norway, 1987-96	36.0 (women) 21.4 (men)	ACR criteria (1987)	(126)
Oslo, 1988-93	36.7 (women) 13.8 (men)	Registry data	(136)
UK, 1991	36 (women) 14 (men)	ACR criteria (1987)	(137)

13. FIGURES

FIGURE 1: AGE & SEX DISTRIBUTION OF PERSON TIME: MONOTHERAPY

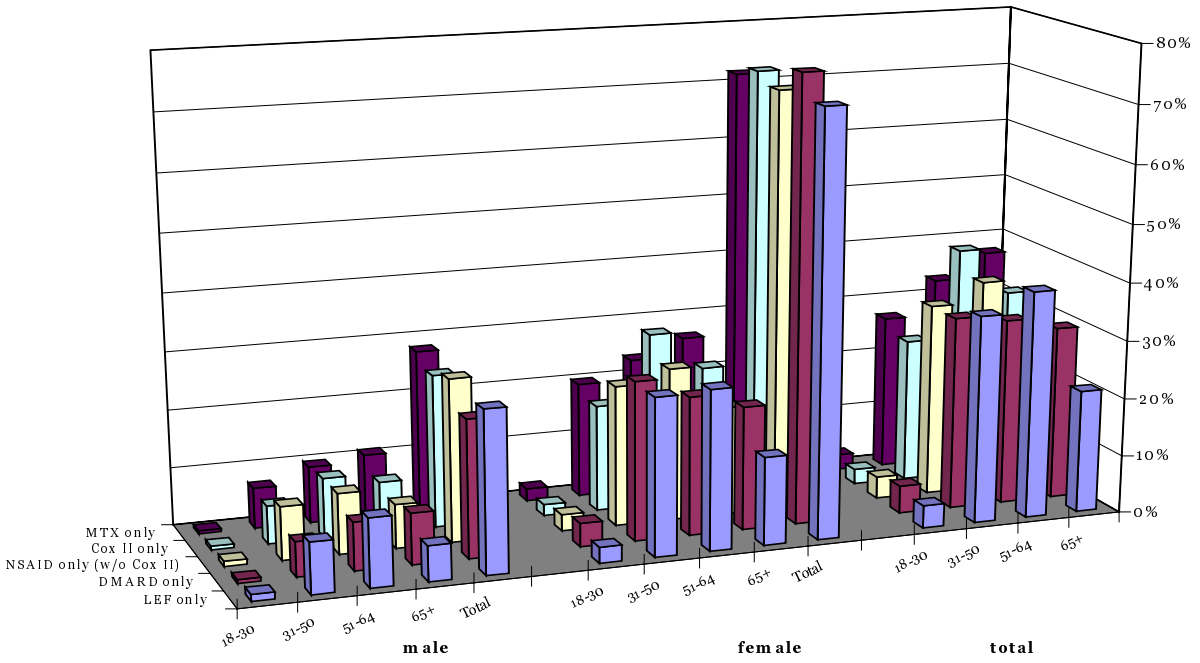


FIGURE 2. MEAN NUMBER OF COMORBIDITIES AT INDEX DATE
*** = statisitcally different result compared to LEF**

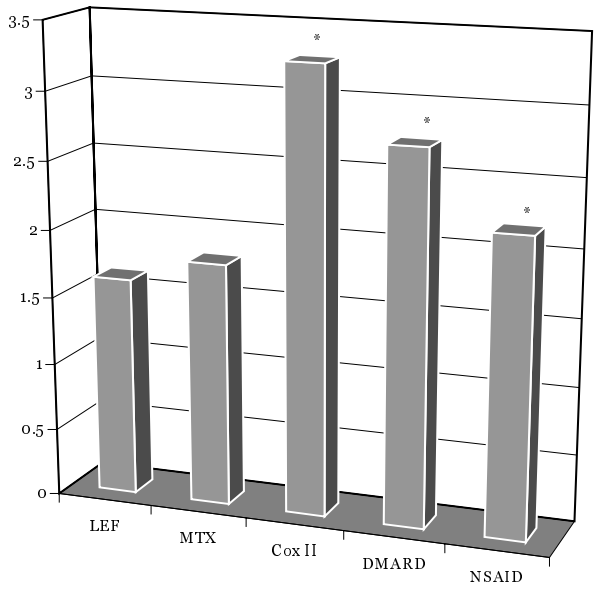
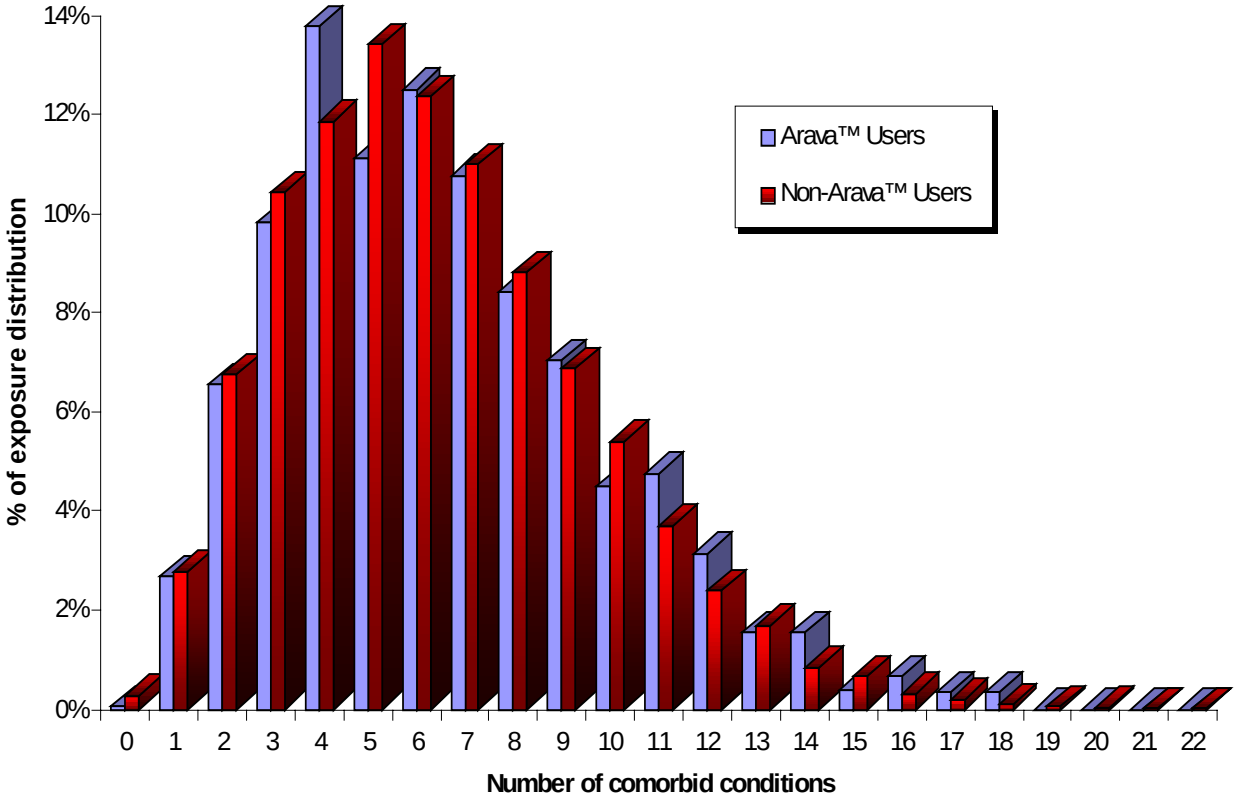


Figure 3. Comparison of Number of Comorbidities Among those Experiencing an Endpoint: Arava users vs. non-Arava users



14. APPENDICES

APPENDIX A: STUDY COHORTS

- **Monotherapy cohorts** (n=3): leflunomide (alone or with NSAID, cox-2), DMARD (alone or with NSAID, cox-2), methotrexate (alone or with NSAID, cox-2)
- **Two-drug therapy combination cohorts** (n=3): leflunomide + DMARD (with or without NSAID, cox-2), leflunomide + methotrexate (with or without NSAID, cox-2), DMARD + methotrexate (with or without NSAID, cox-2)

leflunomide:

leflunomide (leflunomide) _66-28-00-50-00

Methotrexate

Methotrexate Sodium For Inj 1 GM _21-30-00-50-10-21-50
Methotrexate Sodium For Inj 20 MG _21-30-00-50-10-21-05
Methotrexate Sodium Inj 25 MG/ML _21-30-00-50-10-20-30
Methotrexate Tab 2.5 MG _21-30-00-50-00-03-05
Methotrexate Tab 2.5 MG (Antirheumatic) _66-25-00-50-00-03-20

NSAIDs

_66-10-00-07-00 Diclofenac Sodium
_66-10-00-07-10 Diclofenac Potassium
_66-10-00-08-00 Etodolac
_66-10-00-10-10 Fenoprofen Calcium
_66-10-00-12-00 Flurbiprofen
_66-10-00-20-00 Ibuprofen
_66-10-00-30-00 Indomethacin
_66-10-00-30-10 Indomethacin Sodium
_66-10-00-35-00 Ketoprofen
_66-10-00-37-10 Ketorolac Tromethamine
_66-10-00-40-10 Meclofenamate Sodium
_66-10-00-50-00 Mefenamic Acid
_66-10-00-52-00 Meloxicam
_66-10-00-55-00 Nabumetone
_66-10-00-60-00 Naproxen
_66-10-00-60-10 Naproxen Sodium
_66-10-00-65-00 Oxaprozin
_66-10-00-70-00 Piroxicam
_66-10-00-80-00 Sulindac
_66-10-00-90-10 Tolmetin Sodium
_66-10-10-10-00 Phenylbutazone
_66-10-99-02-20 Diclofenac w/ Misoprostol

COX-2 INHIBITORS

_66-10-05-25-00 Celecoxib
_66-10-05-65-00 Rofecoxib

DMARDS

Gold Compunds
_66-20-00-10-00 Auranofin
_66-20-00-20-00 Aurothioglucose
_66-20-00-30-00 Gold Sodium Thiomalate

Soluble Tumor Necrosis Factor receptor antagonist
_66-29-00-30-00 Etanercept (Enbrel)

Anti-TNF antibody
Infliximab (Remicade)

Antimalarials
_13-00-00-20-10 hydroxychloroquine
Chloroquine phosphate, Chloroquine sulphate

Antibiotics
- Minocycline

Chelating agents
_99-20-00-30-00 penicillamine

Sulfasalazine
_52-50-00-60-00 sulfasalazine

Steroids
_22-10 glucocorticoids

Cytotoxics
_21-10-10-10-00 chlorambucil
_21-10-10-20-00-03 cylcophosphamide

cyclosporine

APPENDIX B: ENDPOINT DEFINITIONS

Hepatic Events

Acute or Subacute Liver Necrosis (ICD-9CM : 570)
Cirrhosis of liver without mention of alcohol (ICD-9CM 571.5)
Biliary Cirrhosis (ICD-9CM : 571.6)
Hepatic Coma (ICD-9CM : 572.2)
Hepatitis, Noninfectious Toxic (ICD-9CM : 573.3)
Unspecified chronic liver disease without mention of alcohol (ICD-9CM : 571.9)
Other Specified Liver Disorder (ICD-9CM : 573.8)
Unspecified Liver Disorder (ICD-9CM : 573.9)

Elevation in Enzymes

790.4 SGOT
790.4 SGPT
790.4 transaminase
790.5 acid phosphatase
790.5 alkaline phosphatase

Aplastic Anemia, Pancytopenia

284.0 Constitutional aplastic anemia
Aplasia, (pure) red cell:
congenital
of infants
primary
Blackfan-Diamond syndrome
Familial hypoplastic anemia
Fanconi's anemia
Pancytopenia with malformations

284.8 Other specified aplastic anemias
Aplastic anemia (due to):
chronic systemic disease
drugs
infection
radiation
toxic (paralytic)
Pancytopenia (acquired)
Red cell aplasia (acquired) (adult) (pure) (with thymoma)

284.9 Aplastic anemia, unspecified Anemia:
aplastic (idiopathic) NOS
aregenerative
hypoplastic NOS
nonregenerative
refractory
Medullary hypoplasia

Severe Skin Reactions

695.1 Erythema multiforme
Erythema iris
Herpes iris
Lyell's syndrome
Scalded skin syndrome
Stevens-Johnson syndrome
Toxic epidermal necrolysis

Hypertension

401Essential hypertension
[Exclude:

eye (362.11)
pulmonary hypertension (416.0-416.9)
that involving vessels of:
brain (430-438)]

401.0 Malignant
401.1 Benign
401.9 Unspecified Elevated blood pressure

796.2 elevated blood pressure without diagnosis of hypertension

Pneumonitis

Pneumonitis (acute) (primary)
specified type NEC 495.8
allergic 495.9
hypersensitivity 495.9
chemical 506.0
cholesterol 516.8
lymphoid 516.8
lymphoid, interstitial 516.8
eosinophilic 518.3
515 Postinflammatory pulmonary fibrosis
Cirrhosis of lung chronic or unspecified
Fibrosis of lung (atrophic) (confluent) (massive) (perialveolar) (peribronchial) chronic or unspecified
Induration of lung chronic or unspecified

Pancreatitis

577.0 **Acute pancreatitis**
Abscess of pancreas
Necrosis of pancreas:
acute
infective
Pancreatitis:
NOS
acute (recurrent)
apoplectic
hemorrhagic
subacute
suppurative

790.5 **Elevated amylase**

GI bleeding

Site-specific codes

531.0 Gastric ulcer with hemorrhage
531.1 Gastric ulcer acute with perforation
531.2 Gastric ulcer acute with hemmorhage or perforation
531.4 Gastric ulcer chronic/unspecified with hemorrhage
531.5 Gastric ulcer chronic/unspecified with perforation
531.6 Gastric ulcer chronic/unspecified with hemorrhage or perforation
532.0 Duodenal ulcer acute with hemorrhage
532.1 Duodenal ulcer acute with perforation
532.2 Duodenal ulcer acute with hemorrhage or perforation
532.4 Duodenal ulcer chronic/unspecified with hemorrhage
532.5 Duodenal ulcer chronic/unspecified with perforation
532.6 Duodenal ulcer chronic/unspecified with hemorrhage or perforation
534.0 Gastrojejunal ulcer acute with hemorrhage
534.1 Gastrojejunal ulcer acute with perforation
534.2 Gastrojejunal ulcer acute with hemorrhage or perforation
534.4 Gastrojejunal ulcer chronic/unspecified with hemorrhage
534.5 Gastrojejunal ulcer chronic/unspecific with perforation
534.6 Gastrojejunal ulcer chronic/unspecified with hemorrhage or perforation

Lesion-specific codes

533.0 Peptic ulcer acute with hemorrhage
533.1 Peptic ulcer acute with perforation
533.2 Peptic ulcer acute with hemorrhage or perforation
533.4 Peptic ulcer chronic/unspecified with hemorrhage
533.5 Peptic ulcer chronic/unspecified with perforation
533.6 Peptic ulcer chronic/unspecified with hemorrhage or perforation

Nonspecific codes

578.0 Hematemesis
578.1 Melena
578.9 Hemorrhage of the intestinal tract unspecified

Respiratory Tract Infections: Upper and Bronchitis

465.0 Acute laryngopharyngitis
465.8 Other multiple sites: Multiple URI
465.9 Unspecified site: Acute URI NOS, Upper respiratory infection (acute)
466.0 Acute bronchitis: Bronchitis, acute or subacute:
487.x: Influenza
487.0 With pneumonia
487.1 With other respiratory manifestations
490 Bronchitis, not specified as acute or chronic
490 Bronchitis, not specified as acute or chronic
519.8 Infection respiratory NOS

APPENDIX C: HALF-LIVES OF STUDY MEDICATIONS USED TO DETERMINE PERSON-TIME EXPOSURE

Medication	Upper bound of elimination half-life (in days)	5x elimination half-life (days)
ARAVA™:		
Arava (leflunomide) _66-28-00-50-00	14.0	70 (used 60 in analysis)
Methotrexate		
Methotrexate Sodium For Inj 1 GM _21-30-00-50-10-21-50	0.6	3.1
Methotrexate Sodium For Inj 20 MG _21-30-00-50-10-21-05	0.6	3.1
Methotrexate Sodium Inj 25 MG/ML _21-30-00-50-10-20-30	0.6	3.1
Methotrexate Tab 2.5 MG _21-30-00-50-00-03-05	5.0	25.0
Methotrexate Tab 2.5 MG (Antirheumatic) _66-25-00-50-00-03-20	5.0	25.0
NSAIDs		
_66-10-00-07-00 Diclofenac Sodium	0.1	0.4
_66-10-00-07-10 Diclofenac Potassium	0.1	0.4
_66-10-00-08-00 Etodolac	0.1	0.4
_66-10-00-10-10 Fenoprofen Calcium	0.1	0.6
_66-10-00-12-00 Flurbiprofen	0.2	0.8
_66-10-00-20-00 Ibuprofen	0.2	0.8
_66-10-00-30-00 Indomethacin	0.5	2.3
_66-10-00-30-10 Indomethacin Sodium	0.5	2.3
_66-10-00-35-00 Ketoprofen	0.2	1.0
_66-10-00-37-10 Ketorolac Tromethamine	0.3	1.3
_66-10-00-40-10 Meclofenamate Sodium	0.2	1.0
_66-10-00-50-00 Mefenamic Acid	0.1	0.4
_66-10-00-52-00 Meloxicam	0.8	4.2
_66-10-00-55-00 Nabumetone	1.5	7.3
_66-10-00-60-00 Naproxen	0.5	2.5
_66-10-00-60-10 Naproxen Sodium	0.5	2.5
_66-10-00-65-00 Oxaprozin	2.1	10.4
_66-10-00-70-00 Piroxicam	2.1	10.4
_66-10-00-80-00 Sulindac	0.7	3.4
_66-10-00-90-10 Tolmetin Sodium	0.2	1.0
_66-10-10-10-00 Phenylbutazone	1.0	5.0
_66-10-99-02-20 Diclofenac w/ Misoprostol	0.1	0.4
COX-2 INHIBITORS		
_66-10-05-25-00 Celecoxib	0.5	2.7
_66-10-05-65-00 Rofecoxib	0.7	3.5
DMARDS		
Gold Compunds		0.0
_66-20-00-10-00 Auranofin	26.0	130.0
_66-20-00-20-00 Aurothioglucose	27.0	135.0
_66-20-00-30-00 Gold Sodium Thiomalate		0.0
Soluble Tumor Necrosis Factor receptor antagonist		

_66-29-00-30-00 Etanercept (Enbrel)	4.8	24.0
Anti-TNF antibody		
Infliximab (Remicade)	9.0	45.0
Antimalarials		
_13-00-00-20-10 hydroxychloroquine	14.0	70.0
Chloroquine phosphate, Chloroquine sulphate	5.0	25.0
Antibiotics		
- Minocycline	1.1	5.4
Chelating agents		
_99-20-00-30-00 penicillamine	0.3	1.5
Sulfasalazine		
_52-50-00-60-00 sulfasalazine	0.3	1.5
Steroids		
_22-10 glucocorticoids		0.0
Cytotoxics		
_21-10-10-10-00 chlorambucil	0.1	0.5
_21-10-10-20-00-03 cylcophosphamide	0.5	2.5
Cyclosporine	1.1	5.6

APPENDIX D: CASE VALIDATION/ABSTRACTION FORM

Validation of Cases of Severe Hepatic Events: Arava™ Study

- All patients in the RA cohort with a diagnosis code indicating a hepatic event of interest will be identified. A letter will be sent from the US Quality Assurance (USQA) department at Aetna to the site of care requesting copies of both the office/hospital notes at and around the time of the coded event of interest in addition to any of the following laboratory tests that may have been undertaken: liver biopsy, ultrasound, CT/MRI scan, serum chemistries including liver enzymes, bilirubin and hepatitis titres.
- The office or hospital will receive a financial incentive to respond.
- All responses will be de-identified by USQA
- All responses will be collected centrally by USQA.
- A trained nurse-abstractor will review all returned material and complete as much as possible the attached forms regarding the patient's pre- and post-diagnosis status as outlined in the following abstract forms

PhRMA/FDA/AASLD Drug-Induced Hepatotoxicity White Paper: Postmarketing Considerations. November 2000

If information is available after the cessation of drug (e.g. dechallenge), drug-induced hepatocellular injury is suggested if the decrease of ALT is more than 50% of the excess over the upper limit of normal within 8 days and no additional elevation of ALT within a month. It is suggestive if the decrease is more than 50% within 30 days, and not suggestive if the variation in ALT levels are otherwise.

Causal relationship with a drug can only be excluded when the timing is incompatible -onset of liver damage is before the drug is administered, or when the liver abnormalities are discovered months to years after the drug is withdrawn. Underlying illnesses may mimic drug-induced hepatotoxicity and must be considered in attempting to deduce causality. These include but are not limited to chronic alcoholism (elevated aminotransferase with AST to ALT ratio of more than 2 is suggestive of alcohol liver damage), bacterial infection (elevated alkaline phosphatase or total bilirubin with rare elevations of aminotransferase levels above 5N), right sided or biventricular heart failure, and left sided heart failure or hypotension. Right-sided heart failure may lead to congestion of the liver with subsequent very large elevations of AP and/or transaminases, and/or unconjugated bilirubin in most cases. Left-sided heart failure or hypotension, such as occurs as a result of arrhythmia or myocardial infarction, may result in hypoxia of the liver. In such cases, a rapid rise in aminotransferase levels followed by a rapid return to normal is typical of some cases and sometimes accompanied by a delayed hyperbilirubinemia by 48-72 hours. Viral etiologies must be ruled out since this is the most common cause for hepatocellular injury. Fatty liver of pregnancy (cholestatic injury) and biliary obstruction should likewise be eliminated as a possible etiology by hepatobiliary ultrasound.

Table 1. Definitions and Types of Liver Inquiry (Strawman)

Liver Injury	Hepatocellular	Cholestatic	Mixed
>2-3xULN of ALT (SGPT) OR	>2-3 x ULN in ALT and nl Alk Phos OR	>2 x ULN in Alkaline Phosphate OR	>2-3 x ULN ALT AND >2 x ULN Alkaline Phosphate OR
>2 x ULN conjugated Bilirubin, OR elevated AST (SGOT), Alkaline Phosphate and Total Bilirubin (one of these must be >2 x ULN)	Ratio of ALT to Alkaline Phosphate ≥ 5	Ratio of ALT to Alkaline Phosphate ≤ 2	Ratio of ALT TO Alkaline Phosphate AND Ratio of ALT and Alkaline Phosphate between 2 and 5

Table 2 Sample Proposed Formats for Collecting Hepatotoxicity Data

LIVER INJURY

Signs of severe injury include a marked elevation of ALT or conjugated bilirubin (CB), PT prolongation, the presence of jaundice in association with hepatocellular injury, or the presence of hepatic encephalopathy.

1 Abbreviated List

Shaded cells: most desired data elements.

1. RELEVANT HISTORY AND CLINICAL CONDITIIONS					
Hepatobiliary disorder	No Yes Specify	Alcohol abuse	No Yes Specify	Risk factor(s) for viral hepatitis	No Yes Specify

2. RELEVANT TESTS/LABORATORY DATA (where ND=not determined)							
a). Lab. Tests	Normal Range	**Before** treatment	**During treatment**		**After cessation of the suspected drug** (enter only lowest levels observed) Drug(s) discontinued: _____		
			Earliest Abnormalities Observed*	Highest level observed	Day 0	First 8 days	Day 9-180
DATE (M/D/Y)							
ALT (SGPT)							
AST (SGOT)							
AP (Alk Phos)							
T. Bili. (TB)							
C. Bili. (CB)							
Protime (PT)							
GGT							
CPK							
b).Serology	**Not done**	**Absent**	**Present**	**Titer**	**Date M/D/Y)**	**c) Liver Biopsy**	
Anti HAV/IgM Anti HBc/gM Anti-HCV Other serology						Not done Done Findings	

Table 2 *(Continued)*

2 Comprehensive List

1. RELEVANT HISTORY AND CLINICAL CONDITIIONS						
	No Yes Specify		No Yes Specify		No Yes Specify	
Hepatobiliary disorder		Right side heart failure		Occupational toxic agent		
Alcohol abuse		Recent hypotension		Intravenous drugs abuse		
Drug allergy		Cancer		Acupuncture		
Auto-immune		Transfusion of blood products		Recent travel to Africa, Asia		

2. RELEVANT TESTS/LABORATORY DATA (where ND=not determined)							
a). Lab. Tests	Normal Range	**Before** treatment	**During treatment**		**After cessation of the suspected drug** (enter only lowest levels observed) *[re lowest – why- makes no sense]* Drug(s) discontinued: _____		
			Earliest abnormalities observed*	Highest level observed	Day 0	First 8 days	Day 9-180
DATE (M/D/Y)							
ALT (SGPT)							
AST (SGOT)							
AP (Alk Phos)							
T. Bili. (TB)							
C. Bili. (CB)							
Protime (PT)							
GGT							
CPK							
b).Serology	**Not done**	**Absent**	**Present**	**Titer**	**Date M/D/Y)**	**c) Liver Biopsy**	
Anti HAV/IgM Anti HBc/IgM Anti-HCV Anti-CMV Igm Anti-nuclear Anti-native DNA Anti-smooth muscle Anti-mitochondria Other serology						Not done Done Findings_____ *[Re not done, done findings – what's this mean?]*	

*Enter only result from the same day

FINAL REPORT

A BI-COHORT STUDY OF THE RISKS OF LEFLUNOMIDE AND OTHER DMARDS IN RHEUMATOID ARTHRITIS

Samy Suissa, Ph.D.
Pierre Ernst, MD MSc
Abbas Kezouh, PhD
Alain Bitton, MD
Marie Hudson, MD

McGill University Health Centre
Montreal, Canada

Submitted to:

Dr Wanju Dai
Global Epidemiology
Aventis Pharmaceuticals
400 Crossings Boulevard
Bridgewater, NJ 08876
Tel: 908.231.3040
Fax: 908.231.2288
E-mail: wanju.dai@aventis.com

Correspondence should be addressed to:

Dr. Samy Suissa
McGill University
Director, Division of Clinical Epidemiology
Royal Victoria Hospital
687 Pine avenue west, Ross 4.29
Montréal, Québec, CANADA
H3A 1A1

Tel: 514-843-1564
Fax: 514-843-1493
E-mail: samy.suissa@clinepi.mcgill.ca

January 20, 2003

SUMMARY

Leflunomide, a new Disease Modifying Anti-Rheumatic Drug (DMARD) introduced in 1998, has been the object of several spontaneous reports of adverse events and clinical cases described in the literature. We report on a study conducted in two large databases of health insurance claims to assess the risk of serious hepatic, dermatologic, hematologic and other adverse outcomes associated with the use of leflunomide and other DMARDS, relative to methotrexate.

We formed a retrospective cohort using data from the *Protocare* and *PharMetrics* claims databases that together encompass 26 million lives. The cohort included subjects with a diagnosis of rheumatoid arthritis who filled a prescription for a DMARD between September 1, 1998 and December 31, 2001. Cohort members were followed from the date of their first DMARD to the occurrence of serious hepatic events, hematologic events, severe skin reactions and pancreatitis, all requiring hospitalisation, as well as pneumonitis, opportunistic infection and septicemia, and lymphoma. The composite endpoint defined as the occurrence of any of the above diagnoses was used. The analysis employed a nested case-control approach with 10 to 100 randomly selected controls per case on their index date. The DMARDS dispensed during the year prior to the index date, including leflunomide, the newer biologic DMARDS and the other DMARDS were compared to monotherapy with methotrexate. Conditional logistic regression was used to estimate the rate ratio of the different endpoints, adjusted for age, gender, non-DMARD use and comorbidity.

The *PharMetrics* and *Protocare* cohorts comprised 33,009 and 8,876 users of a DMARD respectively, in which 463 cases from all causes occurred during follow-up. Overall, the rate ratio of the combined outcome of any adverse event requiring hospitalisation for leflunomide use during the year prior to the index date was 1.1 (95% CI: 0.7-1.5) while it was 1.8 (95% CI: 1.2-2.7) for biological DMARDS and 1.2 (95% CI: 0.9-1.5) for other DMARDS. While current use of leflunomide had no increased risk, past use of this medication appeared to be associated with an increased risk (RR 1.7; 95% CI: 1.0-2.9). The risk of serious hepatic events was increased only with biological DMARDS (RR 5.4: 95% CI: 1.2-24.7), while the risk of serious hematological adverse events was not increased with any DMARD. The risk of serious pancreatitis was doubled with biological DMARDS, but not with any other DMARD, inlcuding leflunomide. The risk of opportunistic infections and septicemia was doubled with biological DMARDS, but not with any other DMARD, including leflunomide. The incidence of severe skin reactions, interstitial pneumonias and lymphomas was too low to allow analyses.

Among patients with rheumatoid arthritis treated with a DMARD, we did not find an excess risk of adverse events with use of leflunomide relative to the use of methotrexate as a single disease modifying therapy. The finding of an increased risk in past users of leflunomide is likely an artifact resulting from early recognition of adverse events, compensated by lower risks for current use. The study had sufficient power to detect two-fold increases in the risk of most adverse events, with the exceptions of serious pancreatitis and hepatitis for which the study could detect rate ratios of 2.5 and 5 respectively.

INTRODUCTION

Leflunomide, approved by the US Food and Drug Administration (FDA) in September 1998, was the first new Disease Modifying Anti-Rheumatic Drug (DMARD) introduced in a decade. It is indicated for adults with active rheumatoid arthritis (RA) to reduce signs and symptoms and to retard structural damage as evidenced by X-ray erosions and joint space narrowing. There have been spontaneous reports to the manufacturer and drug regulators as well as clinical cases described in the literature of adverse events in association with the use of leflunomide.

A recent study evaluated the risks of leflunomide and other DMARDS in a cohort of 40,594 patients with RA drawn from the Aetna-US Healthcare claims database covering 10 million persons (*Post-marketing cohort study of Leflunomide and other DMARDs: A comparative risk analysis, Global Epidemiology, Aventis Pharmaceuticals, March 7, 2002*). The cohort spanned the period from September 1998 through December 2000. The principal comparisons involved exposure to one and two drug combinations with leflunomide monotherapy and leflunomide with methotrexate as the reference groups. The events of interest consisted of serious hepatic events, other hepatic events, hematologic events, severe skin reactions, hypertension, vasculitis and hemolytic anemia, pneumonitis, pancreatitis, gastrointestinal bleeding, respiratory events, and septic arthritis, as well as a composite outcome of any of these events. That study found that the rates of these events with leflunomide exposure were statistically lower or no different than for the reference. The results are limited by the lack of more intricate analyses of the cohort, due to the restricted access to the raw database.

We report on another study conducted in two large databases of health insurance claims to assess the risk of serious hepatic, dermatologic, hematologic and other adverse outcomes associated with the use of leflunomide and other DMARDS, relative to methotrexate.

METHODS

Study Design and Data Source

We formed a retrospective cohort, based on two sources of data, to evaluate these risks. The first data source is a subset of the *Protocare* longitudinal health benefit claims database that combines data from Medicaid, Medicare, private health maintenance organizations (HMO) and preferred provider organizations (PPO). This proprietary database encompasses over 10 million lives and has been in existence since 1991. The second source of data is the *PharMetrics* Integrated Outcomes Database. It consists of standardized information on claims data from over 40 different managed care organizations and encompasses more than 16 million lives. For the present study, the two datasets were limited to claims with at least one occurrence of a diagnosis of rheumatoid arthritis (ICD-9: 714) between January 1, 1998 and December 31, 2001. These databases do not permit access to the medical records so as to protect patient confidentiality.

Because of the complexity in the patterns of drugs used to treat RA, the risks were estimated using a nested case-control approach. This allows one to deal effectively with multiple drug use and varying durations of use.

Cohort Definition

Cohort entry was defined for both cohorts by the date of the first prescription for a DMARD after September 1, 1998, the launching date of leflunomide in the US. The DMARDs include leflunomide, methotrexate, gold compounds, anti-tumor necrosis factor alpha agents (anti-TNF), antimalarials, minocycline, chelating agents, sulfasalazine and cytotoxics. All subjects were followed from the date of the first prescription until the earliest of: the date of termination of enrollment in the health plan, the date of death, the end of the study period (December 31, 2001) or the date of the clinical outcome of interest. Subjects had to be eighteen years or older at cohort entry. Subjects with less than three months of eligibility in the health insurance plan prior to cohort entry were excluded. In addition, subjects with the outcome of interest during the three-month period prior to cohort entry were excluded.

Outcome Events

Outcome events were identified from inpatient and outpatient encounters, using specific ICD-9 codes (see Appendix A). The events under study include serious hepatic events (hepatic necrosis, cirrhosis, hepatic coma, and hepatitis), hematologic events (aplastic anemia, agranulocytosis, pancytopenia), severe skin reactions (erythema multiformae, Stevens-Johnson Syndrome, toxic epidermal necrolysis), and pancreatitis all requiring hospitalisation, as well as pneumonitis, opportunistic infection and septicemia, and lymphoma. We also evaluated the same endpoints without the requirement for hospitalisation (expanded definition).

Because of the rarity of some of these outcomes, a composite endpoint, defined as the occurrence of ***any*** of the above diagnoses, was created. For subjects with more than one endpoint, the first occurrence during follow-up was used.

Nested case-control design

We used a nested case-control design within the cohorts. This approach allows us to address the complex patterns of drug exposure with insignificant loss of power. For each case identified in the cohorts, we randomly selected 10 controls from the cohort, after matching on the date of cohort entry and ensuring that they were at risk on the day of the event of the case. That date was designated the index date. For the events with few cases (less than 100), we increased the number of controls to 100 per case.

Exposure Measurement

All drugs received during follow-up, including DMARDS and other non-DMARD RA drugs, were identified from dispensed prescription data. The type, date of filling, and the duration of each prescription dispensed at the time of cohort entry were obtained from the databases.

For the purposes of comparison, the DMARDS were divided into four groups: leflunomide, the newer biologic DMARDS (TNF receptor antagonists: inflixmab and etanercept), the other DMARDS (gold compounds, antimalarials, minocycline, chelating agents, sulfasalazine and cytotoxics; these include auranofin, aurothioglucose, gold sodium thiomalate, hydroxychloroquine, hydroxychloroquine sulfate, minocycline, penicillamine, sulfasalazine, chlorambucil, cyclophosphamide and cyclosporine) and methotrexate (including methotrexate sodium). Methotrexate was used as the reference drug in all comparisons. The other non-DMARD anti-RA drugs, namely glucocorticoids,

non-steroidal anti-inflammatory drugs (NSAIDS) and COX-2 inhibitors, were not used as exposure but rather as covariates.

Covariate information

Age, gender and the source of data (Protocare or Pharmetrics) were used as basic covariates that define the study population. The assessment of comorbid conditions was based on diagnoses made during the observation period. These included cardiovascular disease (ICD-9: 391-400, 402-404, 410-429, 430-453), respiratory illness (ICD-9: 480-519), diabetes (ICD-9: 250), hypertension (ICD-9: 401), hypercholesterolemia (ICD-9: 272.0), cancer(ICD-9: 140-208, 230-239), gastrointestinal conditions (ICD-9: 530-537,555-558), vasculitis (ICD-9: 446.20,446.29,273.2,287,0) and CNS conditions (ICD-9: 320-389). As mentioned above, non-DMARD drugs used for symptomatic relief, namely glucocorticoids, NSAIDS and COX-2 inhibitors, were also used as covariates to control for disease severity.

Data analysis

Total person-time of follow-up in the 2 cohorts was cumulated to estimate the rate of adverse events for each endpoint, including the composite endpoint, under study. Conditional logistic regression was used with the nested case-control samples to estimate the rate ratio of the different endpoints, including the composite, for any use of leflunomide, newer DMARDS and other DMARDS, all relative to methotrexate monotherapy, during the year prior to the index date. Non-use of any DMARD during the one-year period was accounted for in the analysis to maintain the same reference group across comparisons. Leflunomide exposure was further redefined in two ways. First, current use of leflunomide was defined by the last prescription prior to the index date being dispensed within 90 days of the index date, while any other use during the year prior to the index date was designated as past use. Second, the use of leflunomide during the year prior to the index date was separated as monotherapy or multitherapy if other DMARDs were dispensed at any time during that year. All analyses were adjusted for the concurrent use of other DMARDS, the non-DMARDS, namely glucocorticoids, NSAIDS and COX-2 inhibitors, as well as age, gender and co-morbidity.

RESULTS

There were 96,738 subjects in the *PharMetrics* database and 32,063 in the *Protocare* database with at least one occurrence of the diagnosis of rheumatoid arthritis between January 1, 1998 and December 31, 2001. After excluding subjects who were not dispensed a DMARD, who had less than three months of eligibility in the health insurance plan prior to cohort entry, or with outcome of interest prior to cohort entry, the *PharMetrics* cohort comprised 33,009 subjects who received a DMARD after September 1, 1998, while the *Protocare* cohort had 8,876 subjects. The characteristics of the subjects at cohort entry are displayed in Table 1 for both cohorts. Subjects from the Protocare cohort were 10 years older than those from the Pharmetrics cohort.

The *PharMetrics* cohort was followed for a total of 39,286 person-years, while the *Protocare* cohort had 12,029 person-years of follow-up. There were 463 cases of serious adverse events from all causes in the two combined cohorts, 295 in Pharmetrics and 168 in Protocare. Table 2 shows that the rate of any such adverse event was 75 per 10,000

per year in the Pharmetrics cohort and 140 per 10,000 per year in the older Protocare cohort. Rates are also given for specific events. Of note is the small number of severe skin reactions, pneumonitis and lymphomas.

Table 3 provides descriptive information for these cases and their respective controls in both cohorts. Overall, the cases in the Protocare cohort are more than 10 years older than in the Pharmetrics cohort. Follow-up in the Protocare cohort was also longer, 371 days compared to 302 days in the Pharmetrics cohort. The majority of subjects with rheumatoid arthritis were women. A significant proportion of patients had been dispensed glucocorticoids during the year prior to the index date and this was more likely to have occurred among cases than controls in both cohorts. Comorbidity was common and more so among subjects in the Protocare cohort, who were older, and more common in case patients than control patients in both cohorts. The principal comorbidities during the year prior to the index date were cardiovascular diseases, hypertension, CNS complaints, respiratory diseases and diabetes.

Table 4 presents adjusted rate ratios of the combined outcome of any adverse event requiring hospitalisation for disease modifying anti-rheumatoid arthritis drugs compared with the use of methotrexate as the only disease modifying drug. Overall, in the Pharmetrics cohort there was an increase in the risk of any such adverse event for biological DMARDS (RR 1.8). There was no statistically significant increase in the risk of all adverse events combined in either of the cohorts with leflunomide. An exception was with the past use of this medication, as measured by use during the 9-month period preceding the last 90 days prior to the index date. This excess risk with past use of leflunomide was present in both databases.

When examining the risk of serious hepatic events requiring hospitalisation, the number of cases was low so that 100 controls per case had to be selected. In the Pharmetrics cohort, none of the 11 cases of these hepatic events were exposed to leflunomide and only 2 of the 14 in the Protocare cohort (Table 5). When combining the two cohorts, there is a suggestion of an increased risk of hepatic events requiring hospitalisation with the use of biological DMARDS (RR 5.4; 95% CI: 1.2-24.7) and possibly with the other DMARDS (RR 2.3) as compared to the risk for patients receiving methotrexate as the only disease modifying anti-rheumatoid arthritis drug.

When addressing the risk of hematological adverse events requiring hospitalisation (Table 6), the numbers of cases were relatively small (88 and 50 cases in the Pharmetrics and Protocare cohorts, respectively) and therefore required 100 controls per case. Considering the cohorts together, all rate ratios were below 1.0 for leflunomide. There was also no excess risk demonstrable for biological DMARDS or other DMARDS .

In examining the risk of pancreatic events requiring hospitalisation (Table 7), here again the limited number of cases (46 and 38 cases in the Pharmetrics and Protocare cohorts, respectively) justified the use of 100 controls per case. Past use of leflunomide appears associated, although not significantly so, with an increased risk compared with the use Methotrexate as the only disease modifying agent. An increase in risk of similar size was seen with biological DMARDS.

For the risk of opportunistic infections and septicemia requiring hospitalisation (Table 8) there was no statistically significant increase in risk for any or past use of leflunomide when the cohorts were combined. The use of biological DMARDS was associated with a two-fold increase in risk of opportunistic infections and septicemia requiring hospitalisation. Given that severe opportunistic infections and septicemia would

be expected to result in hospitalisation, we did not examine such events in the absence of a hospitalisation.

The incidence of severe skin reactions was extremely small, with only 3 cases requiring hospitalisation, none of which used leflunomide, so that no analyses could be carried out (Table 9). Interstitial pneumonias (pneumonitis) requiring hospitalisation occurred in insufficient numbers (12 cases overall) to allow an analysis of the risk in association with use of disease modifying medications, although one case was exposed to leflunomide (Table 10). Similarly, among the 5 lymphoma cases, none occurred among subjects on leflunomide, while too few cases were seen amongst patients prescribed methotrexate only, biological DMARDS, or other DMARDS to allow any analyses (Table 11).

Appendix B provides these tables separately for the two cohorts, as well as combined. The similarity of findings in the two cohorts justifies the combined analysis.

Similar findings were observed when examining the risk of these adverse events without requiring the need for hospitalisation (see Appendix C). Only for pancreatic events not requiring hospitalisation (Table C.11) was leflunomide associated with a 70% increase in risk, slightly more marked with monotherapy and past use. For this same outcome, there was an approximately 50% increase in risk for biological DMARDS and other DMARDS with only the latter achieving statistical significance when combining the two cohorts.

DISCUSSION

In two large cohorts of patients with rheumatoid arthritis treated with a DMARD, we did not find an excess risk of adverse events among users of leflunomide, particularly the current users, relative to users of methotrexate as monotherapy

When examining specific adverse events, the number of events where hospitalisation occurred was too small to produce informative analyses for severe skin reactions, interstitial pneumonias and lymphomas. Except for one case of interstitial pneumonia, however, no cases had been exposed to leflunomide. For hepatic and hematological events, pancreatitis and opportunistic infections and septicemia requiring hospitalisation, the number of cases varied between 25 and 138 cases. By increasing the number of controls per case, we were able to increase the power and obtain stable risk estimates. For hepatic adverse events and opportunistic infections and septicemia requiring hospitalisation, no risk was found with leflunomide. For hematological events and pancreatitis, there was a small increase in risk with leflunomide, although the risk was mostly limited to past users.

The finding of a 70% increase in the risk of all adverse events combined with past use of leflunomide, mostly observed for hematological events and pancreatitis, is likely an artifact. We believe it most probably represents cessation of the drug by patients or their physician because of approaching adverse events that were recognized. Even for very acute events, note that while past use is defined by the date of the last drug being dispensed more than 90 days before the index date, its use could have continued into that 90-day period and stopped close to the index date. Moreover, the increase observed with past use is compensated by a rate ratio for current use lower than unity. For these reasons, the evaluation of the risk using the one-year period prior to the index date is a more reliable approach that is less likely to be influenced by such actions. Alternatively, of course, this higher rate with past use with borderline statistical significance (RR 1.7; 95% CI: 1.0-2.9) could also simply be due to random error.

We found an 80% increase in the risk of all adverse events requiring hospitalisation associated with the use of biological DMARDs, although this risk was attenuated when the case definition did not require hospitalisation. For hepatic and hematological events, pancreatitis, opportunistic infections and septicemia, we found an increase in risk with biological DMARDS. This small but systematic increase in the risk of all these events was not the object of the current study but requires further investigation. In particular, it should be noted that no analyses were planned or conducted for specific patterns of use for biological DMARDS; such as past use or multitherapy. In any case, such analyses would not have been possible for biological DMARDS because of their later introduction on the market and the small number of subjects that were prescribed these medications in this study. Nevertheless, future research should address these adverse effects.

The current study has several limitations. Firstly, the number of certain adverse events was small, so that it was not possible to study events such as severe skin reactions, interstitial pneumonias and lymphomas. We clearly had sufficient power (80%), however, to detect a rate ratio of 1.5 with leflunomide use for the combined outcome of any adverse event requiring hospitalisation and a rate ratio of 1.2 without requiring hospitalisation. The power was also sufficient to detect a rate ratio of 1.5 for hepatitis without requiring hospitalisation and hematological events not requiring hospitalisation. However, for hepatitis requiring hospitalisation, the study only had sufficient power to detect rate ratios of 5 or more. For pancreatitis requiring hospitalisation, rate ratios of 2.5 could be detected from our study. Finally, for hematological events and opportunistic infections and septicemia requiring hospitalisation, as well as for pancreatitis not requiring hospitalisation, rate ratios of 2 could be detected with 80% power. Thus, overall, this study provides high confidence in excluding a doubling of the risk of most adverse events, and particularly the combined outcome, associated with leflunomide use. The only exceptions are pancreatitis and hepatitis both requiring hospitalisation for which the study can only provide assurance for rate ratios of 2.5 and 5 respectively. A strength of the study that serves to validate the results is the use of two independent cohorts and the marked consistency of findings across the two cohorts. In addition, the various populations represented in the cohorts including Medicaid, Medicare, private health maintenance organizations and preferred provider organizations and over 40 different managed care organizations provide further consistency to the findings.

Because of the relatively short duration of follow-up, it was unfeasible to evaluate long-term effects of these drugs. Nevertheless, the cohorts had an average follow-up of around one year and up to three years. Moreover, by extending the follow-up to December 2001, the study included the most recent available data to assess the safety of leflunomide. In this study, we could not verify the validity of the diagnoses used to identify adverse events. The differences in the incidence of these events in the two cohorts (8.9 versus 18.9 per 1000 for Pharmetrics and Protocare respectively) could suggest that the diagnostic criteria used were not uniform in the two cohorts. However, age alone may explain these differences. In fact, a strong element of validation of the diagnoses is the marked uniformity in the results across the two cohorts for all adverse events. A further limitation of our study is the possibility of residual confounding. The associations between adverse events and the various medications used may have been attenuated or increased if physicians prescribed certain of these medications in subset of patients with or without risk factors for these adverse events. For instance, biological DMARDS may have been preferentially prescribed to subjects with known susceptibility for liver disease. We

attempted to reduce this form of confounding by restricting the analyses to cases and controls who did not have the adverse event under study prior to cohort entry. We also adjusted for co-morbidity that could confound these risk estimates.

In conclusion, in this large bi-cohort study, we did not find an excess risk of serious adverse events with the use of leflunomide relative to methotrexate in patients with rheumatoid arthritis treated with a DMARD. The small but systematic increase in risk observed with biological DMARDS requires further investigation.

Table 1

Characteristics of subjects at cohort entry

	Pharmetrics (n=33,009)	**Protocare (n=8,876)**
Follow-up (mean in days)	436	499
Age (mean in years)	49	59
Gender (% male)	24%	24%
DMARD at cohort entry:		
Methotrexate	45%	56%
Leflunomide	7%	6%
Biologic DMARDS	5%	1%
Other DMARDS	43%	37%
Leflunomide use at any time during follow-up	16%	14%

Table 2

Overall rates (per 10,000 per year) of serious adverse events under study for the Pharmetrics and Protocare cohorts separately and combined

	Pharmetrics (39,285.8 person-years)		Protocare (12,029.2 person-years)		Combined (51,315.0 person-years)	
	Number	**Rate**	**Number**	**Rate**	**Number**	**Rate**
Any event	295	75.09	168	139.66	463	90.23
Hepatic	11	2.80	14	11.64	25	4.87
Hematologic	88	22.40	50	41.57	138	26.89
Pancreatic	46	11.71	38	31.59	84	16.37
Opportunistic infections and septicemia	153	38.95	62	51.54	215	41.90
Severe skin reactions	3	0.76	0	0.00	3	0.58
Pneumonitis	3	0.76	9	7.48	12	2.34
Lymphoma	3	0.76	2	1.66	5	0.97

Table 3

Comparison of cases of any serious adverse event and controls on characteristics, concurrent other drug use and co-morbidity from the Pharmetrics and Protocare cohorts

	Pharmetrics		Protocare	
	Cases	**Controls**	**Cases**	**Controls**
Number	295	2950	168	1680
Age	53 ± 12	50 ± 11	64 ± 13	61 ± 14
Follow-up (days)	302 ± 257	302 ± 256	372 ± 248	371 ± 247
Gender (% male)	22%	24%	19%	22%
Other RA drugs				
NSAIDs	29%	39%	34%	43%
Cox-2 inhibitors	23%	22%	14%	13%
Glucocorticoids	40%	28%	38%	31%
Concurrent diseases				
Cardiovascular	40%	17%	62%	25%
Respiratory	42%	17%	51%	19%
Diabetes	17%	8%	24%	13%
Hypertension	27%	18%	15%	13%
Hypercholesterolemia	9%	11%	21%	18%
Cancer	20%	8%	26%	12%
Gastrointestinal	22%	11%	27%	17%
CNS conditions	49%	37%	45%	34%
Vasculatis	<1%	<1%	<1%	<1%

Table 4

Crude and adjusted rate ratios of any serious adverse event for newer DMARDs relative to methotrexate monotherapy from the combined cohorts

DMARD use in the prior year	Cases (n=463)	Controls (n=4630)	Crude RR	Adjusted* RR	95% CI
Methotrexate only	158	1771	1.0	1.0	Reference
Leflunomide	53	554	1.1	1.1	0.7-1.5
Monotherapy	26	268	1.1	1.0	0.6-1.6
Multitherapy	27	286	1.1	1.1	0.7-1.7
Current use	32	416	0.9	0.8	0.6-1.3
Past use	21	138	1.7	1.7	1.0-2.9
Biologic DMARDS	37	298	1.4	1.8	1.2-2.7
Other DMARDS	184	1729	1.2	1.2	0.9-1.5

* Adjusted for age, gender, cohort, non use of DMARDs in the year prior to the index date, use of NSAIDS, COX-2 inhibitors, glucocorticoids and co-morbidity.

Table 5

Crude and adjusted rate ratios of serious hepatic events for newer DMARDs relative to methotrexate monotherapy

DMARD use in the prior year	Cases (n=25)	Controls (n=2500)	Crude RR	Adjusted* RR	95% CI
Methotrexate only	7	989	1.0	1.0	Reference
Leflunomide	2	270	1.1	0.9	0.2-4.9
Monotherapy	0	117	0.0	0.0	ne
Multitherapy	2	153	1.9	1.6	0.3-8.7
Current use	0	194	0.0	0.0	ne
Past use	2	76	3.8	2.6	0.4-15.5
Biologic DMARDS	4	128	5.2	5.4	1.2-24.7
Other DMARDS	12	911	1.9	2.3	0.8-6.6

* Adjusted for age, gender, cohort, non use of DMARDs in the year prior to the index date, use of NSAIDS, COX-2 inhibitors, glucocorticoids and co-morbidity.

Table 6

Crude and adjusted rate ratios of serious hematologic events for newer DMARDs relative to methotrexate monotherapy

DMARD use in the prior year	Cases (n=138)	Controls (n=13684)	Crude RR	Adjusted* RR	95% CI
Methotrexate only	62	5250	1.0	1.0	Reference
Leflunomide	17	1624	0.9	0.8	0.5-1.5
Monotherapy	8	785	0.9	0.8	0.3-1.6
Multitherapy	9	839	0.9	0.9	0.4-1.9
Current use	13	1210	0.9	0.9	0.5-1.7
Past use	4	414	0.8	0.7	0.2-1.9
Biologic DMARDS	10	814	1.1	1.2	0.6-2.4
Other DMARDS	40	5059	0.7	0.7	0.5-1.0

* Adjusted for age, gender, cohort, non use of DMARDs in the year prior to the index date, use of NSAIDS, COX-2 inhibitors, glucocorticoids and co-morbidity.

Table 7

Crude and adjusted rate ratios of serious pancreatitis events for newer DMARDs relative to methotrexate monotherapy

DMARD use in the prior year	Cases (n=84)	Controls (n=8394)	Crude RR	Adjusted* RR	Adjusted* 95% CI
Methotrexate only	25	3152	1.0	1.0	Reference
Leflunomide	11	996	1.4	1.5	0.7-3.1
Monotherapy	6	461	1.7	1.7	0.7-4.2
Multitherapy	5	535	1.2	1.3	0.5-3.5
Current use	6	730	1.1	1.1	0.5-2.8
Past use	5	266	2.5	2.4	0.9-6.5
Biologic DMARDS	8	542	2.0	2.2	1.0-5.3
Other DMARDS	31	3089	1.3	1.4	0.8-2.4

* Adjusted for age, gender, cohort, non use of DMARDs in the year prior to the index date, use of NSAIDS, COX-2 inhibitors, glucocorticoids and co-morbidity.

Table 8

Crude and adjusted rate ratios of serious opportunistic infections & septicemia events for newer DMARDs relative to methotrexate monotherapy

DMARD use in the prior year	Cases (n=215)	Controls (n=7729)	Crude RR	Adjusted* RR	95% CI
Methotrexate only	63	3224	1.0	1.0	Reference
Leflunomide	25	888	1.1	0.9	0.6-1.6
Monotherapy	12	452	1.0	0.8	0.4-1.6
Multitherapy	13	436	1.2	1.1	0.6-2.1
Current use	14	638	0.9	0.7	0.4-1.4
Past use	11	250	1.9	1.4	0.7-2.9
Biologic DMARDS	18	197	1.5	2.0	1.1-3.6
Other DMARDS	95	2958	1.3	1.2	0.9-1.7

* Adjusted for age, gender, cohort, non use of DMARDs in the year prior to the index date, use of NSAIDS, COX-2 inhibitors, glucocorticoids and co-morbidity.

Table 9

Frequency of severe skin reactions for newer DMARDs and methotrexate monotherapy (Rate ratios are not estimable)

DMARD use in the prior year	Cases (n=3)	Controls (n=30)
Methotrexate only	0	10
Leflunomide	0	3
Monotherapy	0	2
Multitherapy	0	1
Current use	0	2
Past use	0	1
Biologic DMARDS	0	2
Other DMARDS	3	15

Table 10

Frequency of pneumonitis
for newer DMARDs and methotrexate monotherapy
(Rate ratios are not estimable)

DMARD use in the prior year	Cases (n=12)	Controls (n=120)
Methotrexate only	4	52
Leflunomide	1	13
Monotherapy	1	4
Multitherapy	0	9
Current use	1	7
Past use	0	6
Biologic DMARDS	0	4
Other DMARDS	6	40

Table 11

Frequency of lymphoma
for newer DMARDs and methotrexate monotherapy
(Rate ratios are not estimable)

DMARD use in the prior year	Cases (n=5)	Controls (n=50)
Methotrexate only	0	14
Leflunomide	0	8
Monotherapy	0	3
Multitherapy	0	5
Current use	0	4
Past use	0	4
Biologic DMARDS	1	4
Other DMARDS	3	18

Appendix A

Endpoint Definitions

Hepatic Events,requiring hospitalization

Acute or Subacute Liver Necrosis (ICD-9CM : 570)
Cirrhosis of liver without mention of alcohol (ICD-9CM 571.5)
Hepatitis, Noninfectious Toxic (ICD-9CM : 573.3)
Hepatic Coma (ICD-9CM : 572.2)

Hematologic,requiring hospitalization

284.8 Other specified aplastic anemias
Aplastic anemia (due to):
chronic systemic disease
drugs
infection
radiation
toxic (paralytic)
Pancytopenia (acquired)
Red cell aplasia (acquired) (adult) (pure) (with thymoma)

284.9 Aplastic anemia, unspecified Anemia:
aplastic (idiopathic) NOS
aregenerative
hypoplastic NOS
nonregenerative
refractory
Medullary hypoplasia

287.4 Secondary thrombocytopenia
Posttransfusion purpura
Thrombocytopenia due to:
Dilutional
Drugs
Extracorporeal circulation of blood
Platelet alloimmuinzation

288.0 Agranulocytosis

Severe Skin Reactions,requiring hospitalization

695.1 Erythema multiforme
Erythema iris
Herpes iris
Lyell's syndrome
Scalded skin syndrome
Stevens-Johnson syndrome
Toxic epidermal necrolysis

Hypertension,requiring hospitalization

401.0 Malignant Essential hypertension
401.9 Unspecified Elevated blood pressure

Vasculitis

446.20 Hypersensitivity angiitis
446.29 Other specified hypersensitivity angiitis
273.2 Other paraproteinemias: cryglobulinemic purpura or vasculitis
287.0 Allergic purpura

Pneumonitis

495.9 Unspecified allergic alveolitis and pneumonitis
515 Post-inflammatory pulmonary fibrosis
516.8 Other specified alveolar and parietoalveolar pneumonopathies

in conjunction with:
32.28 lung biopsy (open)
32.37 lung biopsy (closed)

Pancreatitis,requiring hospitalization

577.0 Acute pancreatitis
Abscess of pancreas
Necrosis of pancreas:
acute
infective
Pancreatitis:
NOS
acute (recurrent)
apoplectic
hemorrhagic
subacute
suppurative

Lymphoma

202 Other malignant neoplasms of lymphoid and histiocytic tissue

Opportunistic Infections & Septicemia

010-018 tuberculosis
031 diseases due to mycobacteria
038 septicemia
136.3 pneumocystosis

APPENDIX B

COMPARATIVE RESULTS BY COHORT AND COMBINED

Table B.1

Crude and adjusted rate ratios of any serious adverse event for newer DMARDs relative to methotrexate monotherapy

	PHARMETRICS				PROTOCARE				COMBINED			
DMARD use in the prior year	Cases (n=295)	Controls (n=2950)	Crude RR	Adjusted* RR (95% CI)	Cases (n=168)	Controls (n=1680)	Crude RR	Adjusted* RR (95% CI)	Cases (n=463)	Controls (n=4630)	Crude RR	Adjusted* RR (95% CI)
Methotrexate only	77	975	1.0	Reference	81	796	1.0	Reference	158	1771	1.0	Reference
Leflunomide	35	384	1.2	1.1 (0.7-1.7)	18	170	1.0	1.0 (0.6-1.9)	53	554	1.1	1.1 (0.7-1.5)
Monotherapy	19	194	1.2	1.1 (0.7-2.0)	7	74	0.9	0.9 (0.4-2.1)	26	268	1.1	1.0 (0.6-1.6)
Multitherapy	16	190	1.1	1.1 (0.6-2.0)	11	96	1.1	1.2 (0.6-2.5)	27	286	1.1	1.1 (0.7-1.7)
Current use	20	293	0.9	0.8 (0.5-1.4)	12	123	1.0	0.9 (0.5-1.9)	32	416	0.9	0.8 (0.6-1.3)
Past use	15	91	2.1	2.0 (1.1-3.8)	6	47	1.3	1.3 (0.5-3.6)	21	138	1.7	1.7 (1.0-2.9)
Biologic DMARDS	35	286	1.6	1.9 (1.2-3.0)	2	12	1.7	1.6 (0.3-8.6)	37	298	1.4	1.8 (1.2-2.7)
Other DMARDS	127	1148	1.4	1.4 (1.0-1.9)	57	581	1.0	1.0 (0.7-1.5)	184	1729	1.2	1.2 (0.9-1.5)

* Adjusted for age, gender, non use of DMARDs in the year prior to the index date, use of NSAIDS, COX-2 inhibitors, glucocorticoids and co-morbidity.

Table B.2

Crude and adjusted rate ratios of any hepatic event for newer DMARDs relative to methotrexate monotherapy

	PHARMETRICS				PROTOCARE				COMBINED			
DMARD use in the prior year	Cases (n=11)	Controls (n=1100)	Crude RR	Adjusted* RR (95% CI)	Cases (n=14)	Controls (n=1400)	Crude RR	Adjusted* RR (95% CI)	Cases (n=25)	Controls(n=2500)	Crude RR	Adjusted* RR (95% CI)
Methotrexate only	2	364	1.0	Reference	5	625	1.0	Reference	7	989	1.0	Reference
Leflunomide	0	117	0.0	0.0 (ne)	2	153	1.6	1.8 (0.3-11.8)	2	270	1.1	0.9 (0.2-4.9)
Monotherapy	0	56	0.0	0.0 (ne)	0	61	0.0	0.0 (ne)	0	117	0.0	0.0 (ne)
Multitherapy	0	61	0.0	0.0 (ne)	2	92	2.8	3.5(0.5-23.0)	2	153	1.9	1.6(0.3-8.7)
Current use	0	81	0.0	0.0 (ne)	0	113	0.0	0.0 (ne)	0	194	0.0	0.0 (ne)
Past use	0	36	0.0	0.0 (ne)	2	40	6.4	15.0 (2.2-103.6)	2	76	3.8	2.6 (0.4-15.5)
Biologic DMARDS	2	119	3.0	3.6 (0.4-35.5)	2	9	35.0	34.0(2.5-471.3)	4	128	5.2	5.4 (1.2-24.7)
Other DMARDS	7	401	3.2	4.4 (0.8-25.6)	5	510	1.3	1.5 (0.3-6.4)	12	911	1.9	2.3 (0.8-6.6)

* Adjusted for age, gender, non use of DMARDs in the year prior to the index date, use of NSAIDS, COX-2 inhibitors, glucocorticoids and co-morbidity.

Table B.3

Crude and adjusted rate ratios of any hematologic event for newer DMARDs relative to methotrexate monotherapy

	PHARMETRICS				PROTOCARE				COMBINED			
DMARD use in the prior year	Cases (n=88)	Controls (n=8795)	Crude RR	Adjusted* RR (95% CI)	Cases (n=50)	Controls (n=4889)	Crude RR	Adjusted* RR (95% CI)	Cases (n=138)	Controls (n=13684)	Crude RR	Adjusted* RR (95% CI)
Methotrexate only	34	2965	1.0	Reference	28	2285	1.0	Reference	62	5250	1.0	Reference
Leflunomide	11	1096	0.9	0.8(0.4-1.7)	6	528	0.9	0.8 (0.3-2.1)	17	1624	0.9	0.8 (0.5-1.5)
Monotherapy	6	539	1.0	0.9 (0.4-2.2)	2	246	0.7	0.5 (0.1-2.2)	8	785	0.9	0.8 (0.3-1.6)
Multitherapy	5	557	0.8	0.8 (0.3-2.1)	4	282	1.2	1.2 (0.4-3.6)	9	839	0.9	0.9 (0.4-1.9)
Current use	8	815	0.9	0.9 (0.4-1.9)	5	395	1.1	0.9 (0.3-2.5)	13	1210	0.9	0.9 (0.5-1.7)
Past use	3	281	0.9	0.8 (0.2-2.7)	1	133	0.6	0.5 (0.1-4.0)	4	414	0.8	0.7 (0.2-1.9)
Biologic DMARDS	9	769	1.0	1.1 (0.5-2.5)	1	45	1.9	2.0(0.2-18.6)	10	814	1.1	1.2 (0.6-2.4)
Other DMARDS	27	3367	0.7	0.7 (0.4-1.2)	13	1692	0.6	0.6 (0.3-1.2)	40	5059	0.7	0.7 (0.5-1.0)

* Adjusted for age, gender, non use of DMARDs in the year prior to the index date, use of NSAIDS, COX-2 inhibitors, glucocorticoids and co-morbidity.

Table B.4

Crude and adjusted rate ratios of any pancreatitis event for newer DMARDs relative to methotrexate monotherapy

	PHARMETRICS				PROTOCARE				COMBINED			
DMARD use in the prior year	Cases (n=46)	Controls (n=4600)	Crude RR	Adjusted* RR (95% CI)	Cases (n=38)	Controls (n=3794)	Crude RR	Adjusted* RR (95% CI)	Cases (n=84)	Controls (n=8394)	Crude RR	Adjusted* RR (95% CI)
Methotrexate only	8	1440	1.0	Reference	17	1712	1.0	Reference	25	3152	1.0	Reference
Leflunomide	5	634	1.5	1.5 (0.5-4.7)	6	362	1.7	1.7 (0.7-4.7)	11	996	1.4	1.5 (0.7-3.1)
Monotherapy	4	328	2.2	2.2 (0.6-7.5)	2	133	1.5	1.6 (0.3-7.4)	6	461	1.7	1.7 (0.7-4.2)
Multitherapy	1	306	0.6	0.7 (0.1-5.5)	4	229	1.8	1.8 (0.6-5.8)	5	535	1.2	1.3 (0.5-3.5)
Current use	2	470	0.8	0.8 (0.2-3.8)	4	260	1.6	1.7 (0.6-5.5)	6	730	1.1	1.1 (0.5-2.8)
Past use	3	164	3.6	3.9 (1.0-15.6)	2	102	2.0	1.7 (0.4-8.0)	5	266	2.5	2.4 (0.9-6.5)
Biologic DMARDS	8	485	3.2	3.3 (1.2-9.3)	0	57	0.0	0.0(ne)	8	542	2.0	2.2 (1.0-5.3)
Other DMARDS	20	1749	2.1	2.2(0.9-5.0)	11	1340	0.8	0.9 (0.4-2.0)	31	3089	1.3	1.4 (0.8-2.4)

* Adjusted for age, gender, non use of DMARDs in the year prior to the index date, use of NSAIDS, COX-2 inhibitors, glucocorticoids and co-morbidity.

Table B.5

Crude and adjusted rate ratios of any opportunistic infections & septicemia event for newer DMARDs relative to methotrexate monotherapy

	PHARMETRICS				PROTOCARE				COMBINED			
DMARD use in the prior year	Cases (n=153)	Controls (n=1530)	Crude RR	Adjusted* RR (95% CI)	Cases (n=62)	Controls (n=6199)	Crude RR	Adjusted* RR (95% CI)	Cases (n=215)	Controls (n=7729)	Crude RR	Adjusted* RR (95% CI)
Methotrexate only	33	491	1.0	Reference	30	2733	1.0	Reference	63	3224	1.0	Reference
Leflunomide	21	199	1.6	1.3 (0.7-2.5)	4	689	0.5	0.5 (0.2-1.4)	25	888	1.1	0.9 (0.6-1.6)
Monotherapy	10	106	1.4	1.0 (0.5-2.3)	2	346	0.5	0.5 (0.1-2.2)	12	452	1.0	0.8 (0.4-1.6)
Multitherapy	11	93	1.8	1.7 (0.8-3.8)	2	343	0.5	0.5 (0.1-2.0)	13	436	1.2	1.1 (0.6-2.1)
Current use	11	150	1.1	1.0 (0.4-2.1)	3	488	0.6	0.5 (0.2-1.8)	14	638	0.9	0.7 (0.4-1.4)
Past use	10	49	3.2	2.4 (1.0-5.8)	1	201	0.5	0.4 (0.1-3.0)	11	250	1.9	1.4 (0.7-2.9)
Biologic DMARDS	18	139	2.0	2.8 (1.4-5.5)	0	58	0.0	0.0 (ne)	18	197	1.5	2.0 (1.1-3.6)
Other DMARDS	72	627	1.7	1.6(1.0-2.6)	23	2331	0.9	0.9 (0.5-1.5)	95	2958	1.3	1.2 (0.9-1.7)

* Adjusted for age, gender, non use of DMARDs in the year prior to the index date, use of NSAIDS, COX-2 inhibitors, glucocorticoids and co-morbidity.

Table B.6

Frequency of severe skin reactions
for newer DMARDs and methotrexate monotherapy
(Rate ratios not estimable)

	PHARMETRICS		PROTOCARE		COMBINED	
DMARD use in the prior year	Cases (n=3)	Controls (n=30)	Cases (n=0)	Controls (n=0)	Cases (n=3)	Controls (n=30)
Methotrexate only	0	10	0	0	0	10
Leflunomide	0	3	0	0	0	3
Monotherapy	0	2	0	0	0	2
Multitherapy	0	1	0	0	0	1
Current use	0	2	0	0	0	2
Past use	0	1	0	0	0	1
Biologic DMARDS	0	2	0	0	0	2
Other DMARDS	3	15	0	0	3	15

Table B.7

Frequency of pneumonitis
for newer DMARDs and methotrexate monotherapy
(Rate ratios not estimable)

	PHARMETRICS		PROTOCARE		COMBINED	
DMARD use in the prior year	Cases (n=3)	Controls (n=30)	Cases (n=9)	Controls (n=90)	Cases (n=12)	Controls (n=120)
Methotrexate only	1	13	3	39	4	52
Leflunomide	0	4	1	9	1	13
Monotherapy	0	2	1	2	1	4
Multitherapy	0	2	0	7	0	9
Current use	0	3	1	4	1	7
Past use	0	1	0	5	0	6
Biologic DMARDS	0	4	0	0	0	4
Other DMARDS	1	7	5	33	6	40

Table B.8

Frequency of lymphoma
for newer DMARDs and methotrexate monotherapy
(Rate ratios not estimable)

	PHARMETRICS		PROTOCARE		COMBINED	
DMARD use in the prior year	Cases (n=3)	Controls (n=30)	Cases (n=2)	Controls (n=20)	Cases (n=5)	Controls (n=50)
Methotrexate only	0	8	0	6	0	14
Leflunomide	0	5	0	3	0	8
Monotherapy	0	3	0	0	0	3
Multitherapy	0	2	0	3	0	5
Current use	0	2	0	2	0	4
Past use	0	3	0	1	0	4
Biologic DMARDS	1	3	0	1	1	4
Other DMARDS	1	8	2	10	3	18

APPENDIX C

RESULTS OF ADVERSE EVENTS DEFINED
WITHOUT REQUIREMENT OF HOSPITALISATION

(EXPANDED DEFINITION)

Table C.1

Crude and adjusted rate ratios of any adverse event for newer DMARDs relative to methotrexate monotherapy

(EXPANDED DEFINITION)

	PHARMETRICS				PROTOCARE				COMBINED			
DMARD use in the prior year	Cases (n=1118)	Controls (n=11180)	Crude RR	Adjusted* RR (95% CI)	Cases (n=361)	Controls (n=3610)	Crude RR	Adjusted* RR (95% CI)	Cases (n=463)	Controls (n=4630)	Crude RR	Adjusted* RR (95% CI)
Methotrexate only	311	3757	1.0	Reference	167	1677	1.0	Reference	478	5434	1.0	Reference
Leflunomide	171	1566	1.3	1.3 (1.0-1.6)	44	370	1.2	1.2 (0.8-1.7)	215	1936	1.3	1.2 (1.0-1.5)
Monotherapy	85	881	1.2	1.1 (0.9-1.4)	18	182	1.0	0.9 (0.5-1.5)	103	1063	1.1	1.0 (0.8-1.3)
Multitherapy	86	685	1.5	1.5 (1.1-1.9)	26	188	1.4	1.4 (0.9-2.3)	112	873	1.5	1.5 (1.2-1.8)
Current use	123	1271	1.2	1.1 (0.9-1.4)	28	267	1.1	1.0 (0.7-1.6)	151	1538	1.1	1.1 (0.9-1.3)
Past use	48	295	2.0	1.8 (1.3-2.5)	16	103	1.6	1.4 (0.8-2.5)	64	398	1.9	1.7 (1.2-2.2)
Biologic DMARDS	117	998	1.4	1.4 (1.1-1.8)	4	34	1.2	1.5 (0.5-4.6)	121	1032	1.4	1.4 (1.1-1.7)
Other DMARDS	460	4330	1.3	1.2 (1.1-1.4)	125	1304	1.0	1.0 (0.7-1.3)	585	5634	1.2	1.2 (1.0-1.3)

* Adjusted for age, gender, non use of DMARDs in the year prior to the index date, use of NSAIDS, COX-2 inhibitors, glucocorticoids and co-morbidity.

Table C.2

Crude and adjusted rate ratios of any hepatic event for newer DMARDs relative to methotrexate monotherapy

(EXPANDED DEFINITION)

	PHARMETRICS				PROTOCARE				COMBINED			
DMARD use in the prior year	Cases (n=332)	Controls (n=3320)	Crude RR	Adjusted* RR (95% CI)	Cases (n=90)	Controls (n=8886)	Crude RR	Adjusted* RR (95% CI)	Cases (n=422)	Controls (n=12206)	Crude RR	Adjusted* RR (95% CI)
Methotrexate only	93	1041	1.0	Reference	45	4073	1.0	Reference	138	5114	1.0	Reference
Leflunomide	45	499	1.0	1.0 (0.7-1.4)	11	874	1.1	1.0 (0.5-1.9)	56	1373	1.0	1.0 (0.7-1.3)
Monotherapy	23	274	0.9	0.8 (0.5-1.4)	5	425	1.1	0.9 (0.4-2.3)	28	699	0.9	0.8 (0.5-1.3)
Multitherapy	22	225	1.1	1.1 (0.7-1.9)	6	449	1.2	1.0 (0.4-2.5)	28	674	1.1	1.1 (0.7-1.7)
Current use	32	387	0.9	0.9 (0.6-1.4)	6	633	0.8	0.7 (0.3-1.8)	38	1020	0.9	0.9 (0.6-1.3)
Past use	13	112	1.3	1.0 (0.5-2.0)	5	241	1.9	1.5 (0.6-4.0)	18	353	1.4	1.2 (0.7-2.0)
Biologic DMARDS	38	283	1.5	1.6(1.1-2.5)	2	76	2.4	2.5(0.6-10.7)	40	359	1.5	1.6 (1.1-2.4)
Other DMARDS	146	1360	1.2	1.1 (0.9-1.5)	29	3325	0.8	0.7 (0.5-1.2)	175	4685	1.1	1.0 (0.8-1.3)

* Adjusted for age, gender, non use of DMARDs in the year prior to the index date, use of NSAIDS, COX-2 inhibitors, glucocorticoids and co-morbidity.

Table C.3

Crude and adjusted rate ratios of any hematologic event for newer DMARDs relative to methotrexate monotherapy

(EXPANDED DEFINITION)

	PHARMETRICS				PROTOCARE				COMBINED			
DMARD use in the prior year	Cases (n=533)	Controls (n=5330)	Crude RR	Adjusted* RR (95% CI)	Cases (n=155)	Controls (n=1550)	Crude RR	Adjusted* RR (95% CI)	Cases (n=688)	Controls (n=6880)	Crude RR	Adjusted* RR (95% CI)
Methotrexate only	165	1801	1.0	Reference	71	700	1.0	Reference	236	2501	1.0	Reference
Leflunomide	89	724	1.4	1.3 (1.0-1.8)	20	162	1.2	1.2 (0.7-2.2)	109	886	1.3	1.3 (1.0-1.7)
Monotherapy	40	414	1.1	1.0 (0.7-1.5)	6	87	0.7	0.7 (0.3-1.7)	46	501	1.0	1.0 (0.7-1.4)
Multitherapy	49	310	1.8	1.8 (1.2-2.5)	14	75	1.9	1.9 (1.0-3.9)	63	385	1.8	1.8 (1.3-2.4)
Current use	66	561	1.3	1.3 (0.9-1.7)	12	113	1.0	1.1 (0.5-2.1)	78	674	1.2	1.2 (0.9-1.6)
Past use	23	163	1.6	1.5 (1.0-2.5)	8	49	1.6	1.6 (0.7-3.8)	31	212	1.6	1.6 (1.0-2.4)
Biologic DMARDS	50	415	1.3	1.4 (1.0-2.0)	3	31	1.0	1.4 (0.4-4.9)	53	446	1.3	1.4 (1.0-1.9)
Other DMARDS	198	2156	1.0	1.0 (0.8-1.2)	51	556	0.9	0.9 (0.6-1.4)	249	2712	1.0	1.0 (0.8-1.2)

* Adjusted for age, gender, non use of DMARDs in the year prior to the index date, use of NSAIDS, COX-2 inhibitors, glucocorticoids and co-morbidity.

Table C.4

Crude and adjusted* rate ratios of any pancreatics event for newer DMARDs relative to methotrexate monotherapy

(EXPANDED DEFINITION)

	PHARMETRICS				PROTOCARE				COMBINED			
DMARD use in the prior year	Cases (n=110)	Controls (n=1100)	Crude RR	Adjusted* RR (95% CI)	Cases (n=69)	Controls (n=6850)	Crude RR	Adjusted* RR (95% CI)	Cases (n=179)	Controls (n=7950)	Crude RR	Adjusted* RR (95% CI)
Methotrexate only	21	354	1.0	Reference	30	3231	1.0	Reference	51	3585	1.0	Reference
Leflunomide	20	154	2.2	1.9 (1.0-3.8)	10	642	1.7	1.8 (0.9-3.9)	30	796	1.9	1.7 (1.0-2.8)
Monotherapy	13	82	2.7	2.3 (1.1-4.9)	5	290	1.9	1.9 (0.7-5.2)	18	372	2.2	1.9 (1.1-3.5)
Multitherapy	7	72	1.7	1.5 (0.6-3.8)	5	352	1.5	1.7 (0.7-4.6)	12	424	1.5	1.4 (0.7-2.9)
Current use	15	125	2.1	1.9 (0.9-3.9)	7	443	1.7	1.8 (0.8-4.3)	22	568	1.8	1.7 (1.0-2.9)
Past use	5	29	3.0	2.1(0.7-6.4)	3	199	1.6	1.8 (0.5-6.3)	8	228	2.2	1.8 (0.8-4.2)
Biologic DMARDS	12	104	2.0	2.0 (0.9-4.4)	0	76	0.0	0.0 (ne)	12	180	1.5	1.5 (0.8-3.1)
Other DMARDS	49	424	1.9	2.0 (1.2-3.5)	25	2454	1.1	1.2 (0.7-2.1)	74	2878	1.5	1.6 (1.1-2.3)

* Adjusted for age, gender, non use of DMARDs in the year prior to the index date, use of NSAIDS, COX-2 inhibitors, glucocorticoids and co-morbidity.

Table C.5

Crude and adjusted rate ratios of any skin reaction for newer DMARDs relative to methotrexate monotherapy

(EXPANDED DEFINITION)

	PHARMETRICS				PROTOCARE				COMBINED			
DMARD use in the prior year	Cases (n=30)	Controls (n=2930)	Crude RR	Adjusted* RR (95% CI)	Cases (n=1)	Controls (n=100)	Crude RR	Adjusted RR (95% CI)	Cases (n=31)	Controls (n=2940)	Crude RR	Adjusted* RR (95% CI)
Methotrexate only	7	1027	1.0	Reference	0	29			7	1056	1.0	Refence
Leflunomide	4	399	1.5	1.4 (0.4-4.9)	0	11			4	410	1.5	1.3 (0.4-4.7)
Monotherapy	3	229	1.9	1.9 (0.5-7.8)	0	7			3	236	1.9	1.8 (0.4-7.3)
Multitherapy	1	170	0.9	0.8 (0.1-6.4)	0	4			1	174	0.9	0.7 (0.1-6.3)
Current use	4	336	1.8	1.6 (0.5-5.8)	0	6			4	342	1.8	1.6 (0.5-5.7)
Past use	0	63	0.0	0.0 (ne)	0	5			0	68	0.0	0.0 (ne)
Biologic DMARDS	2	255	1.2	0.8 (0.1-4.5)	0	0			2	255	1.2	1.0 (0.2-5.0)
Other DMARDS	14	1167	1.8	1.6 (0.6-4.2)	1	49			15	1216	1.9	1.8 (0.7-4.4)

* Adjusted for age, gender, non use of DMARDs in the year prior to the index date, use of NSAIDS, COX-2 inhibitors, glucocorticoids and co-morbidity.

Global Epidemiology
23 January 2003

OBJECTIVE

As part of a continuing risk management program for leflunomide, the objective of this study was to examine the spontaneous reports and evaluate potential signals of adverse events associated with leflunomide. Two related methods were employed in these studies: proportional reporting analysis and reporting rate analysis.

METHODS -- PRR

The first method used was a proportional reporting ratio (PRR) analysis. PRR analysis is a relatively new method, developed to compare spontaneous reports of suspected adverse reactions of different drugs when the true number of patients exposed to a drug is unknown.[1] It is one of the few quantitative methods available to evaluate spontaneous reports in the absence of denominator data. The other common approach, discussed below, is to obtain prescription or other usage data to calculate reporting rates of AEs, although there are concerns that these data do not reflect actual usage and that the reporting rates will be confused for incidence rates.[2] Further, the spontaneous reports themselves are subject to numerous biases:

[1] Evans SJW, Waller P, Davis S. Proportional reporting ratios: the uses of epidemiological methods for signal generation. *Pharmacoepidemiol Drug Safety* 1998:7(suppl 2):S102.
[2] Evans SJW. Pharmacovigilance: a science or fielding emergencies? *Stat Med* 2000;19:3199-209.

false signals may be generated due to the ambiguity involved in recognizing AEs[3]; only a fraction of AEs are reported and this proportion is difficult to estimate; physician lack of awareness of both the value and the requirements of reporting; the length of time the drug of interest is on the market; the reporting environment; and the effects of channeling, ie, differential prescribing due to severity of disease.

Nonetheless, PRR analysis is a useful statistical tool, widely employed by the Medicines Control Agency (MCA) in the UK. It is calculated using a 2 x 2 table, as follows:

	reaction of interest	all other reactions	
drug of interest	*a*	*b*	M_1
all other drugs	*c*	*d*	M_2
	N_1	N_2	*N*

PRR is calculated in several ways. One way is a relative risk approach, in which the PRR is determined as a/(a + b) divided by c/(c + d). A second way, which results in very similar findings, is the Bayesian Empirical method, in which PRR is calculated as an observed to expected number of events, thus: a/[(M1*N1)/N]. This latter method is the one used in this report. Criteria to interpret the PRR come from several sources and are similar: a minimum of three reported cases are needed; a PRR of at least 3 and an associated X^2 over 5 [ref4] or a PRR of at least 2 and an associated X^2 over 4[ref5] are considered possible signals.

[3] Koch-Weser J, Sellers EM, Zacest R. The ambiguity of adverse drug reactions. *Eur J Clin Pharmacol* 1977;11:75-8.
[4] Wiholm B-E, Olsson S, Moore N, et al. Spontaneous reporting systems outside the US. In: Strom B, editor. *Pharmacoepidemiology*, 3rd edition. John Wiley & Sons, New York, 2000. pp 175-92.

One way to avoid part of the heterodemicity bias[6] (discordant numerators and denominators) that may be present in PRR analysis is to limit the analysis to the same time periods. This approach was used in the current analysis.

Spontaneous report data used are limited in that there is no way to assess the indication for a particular drug, so in the situation where a specific drug is used for more than one condition (eg, as is the case with methotrexate), there is no way of adjusting for potential confounding by indication. Because the PRR is analogous to the proportionate mortality ratio, a commonly used measure in epidemiology, it suffers from a similar weakness: it can show wide fluctuations unrelated to a drug's true adverse event profile according to what is happening with a comparison drug's adverse event profile. It is, therefore, always important to remember that the PRR represents a proportion amongst an array of events, not an actual occurrence rate. Despite the flaws inherent in PRR, it can nevertheless be used to indicate the relative magnitude of a problem with a given drug at a particular time.

The time period for the PRR analysis was 1 September 1998 through 31 June 2002. Reported events for leflunomide were compared with reports for all other drugs in FDA's AERS database. The analysis used software from QED Solutions (Qscan), and the PRRs were generated using the Bayesian Empirical approach (PRRs using the relative risk approach were virtually identical).

---

[5] Stephens MDB. Causality assessment and signal recognition. In: Stephens MDB, Talbot JCC, Routledge PA, editors. *Detection of new adverse drug reactions*, 4th edition. Macmillan, London, 1998. pp 297-318.

[6] Feinstein AR. *Clinical epidemiology: the architecture of clinical research*. WB Saunders, Philadelphia, 1985

Reporting rate analysis was the second method used to generate signals. This method, which might be described as data mining with denominators, employs spontaneous report data as numerators, with all the caveats outlined above, and drug-specific usage data as denominators. The latter were obtained from IMS sales figures, which were converted into person-year exposure using the prescription dose. Again, these rates reflect reporting intensity, not occurrence rates of adverse events, which can only be determined from epidemiologic studies. Reporting rates were calculated for leflunomide, etanercept, infliximab, and methotrexate. The time period covered in this analysis is October 1998 through June 2002. The different events of interest were defined according to specific MedDRA terms as follows:

Hepatic failure - hepatic failure

Interstitial lung disease – interstitial pneumonitis, interstitial lung disease, pneumonitis NOS

Sepsis/TB – sepsis NOS, bacteremia, pulmonary sepsis, neutropenic sepsis, pulmonary tuberculosis, tuberculosis NOS, reactivated tuberculosis

Bullous conditions – erythema multiformae, toxic epidermal necrolysis, Stevens Johnson syndrome

Lymphoma – Hodgkin's disease, lymphoma NEC, non-Hodgkin's b-cell, non-Hodgkin's t-cell, lymphoma unspecified

Hypertension – accelerated hypertension, hypertensive crisis, malignant hypertension NOS, diastolic hypertension, systolic hypertension, hypertension NOS, labile hypertension, aggravated hypertension, essential hypertension

Vasculitis – anti-neutrophil cytoplasmic antibody positive vasculitis, vasculitic rash, leukocytoclastic vasculitis, skin vasculitis NOS, vasculitis NOS, vascular purpura

Pancytopenia – marrow depression and hypoplastic anemias, thrombocytopenia, aggravated thrombocytopenia

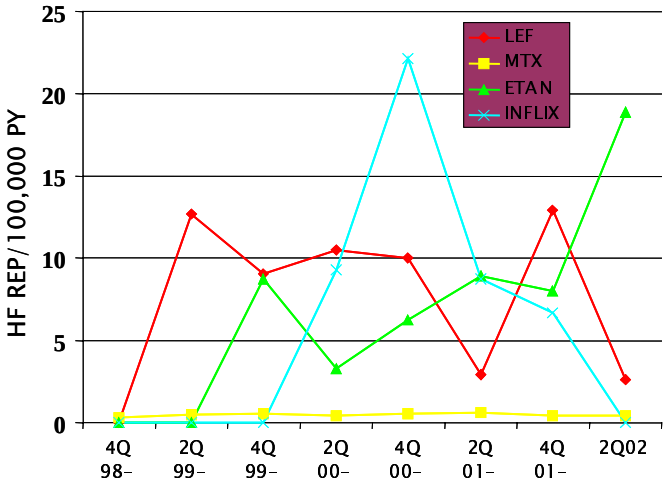
Results of the PRR analysis are shown below for several high level terms incorporating hepatic, hypertensive, hematologic, severe cutaneous, pancreatic, respiratory, vasculitic, and oncologic events. The events of interest were defined by MedDRA terms, which are identified as either System Organ Class (SOC), High Level Group Terms (HLGT), High Level Terms (HLT), or Preferred Terms (PT), levels connoting successively more specificity, along with the number of reports (N).

EVENTS

EVENT, N	PRR	X^2
hepatic and hepatobiliary disorders (HLGT), 243	1.27	13.9
hepatocellular damage and hepatitis NEC (HLT), 68	1.57	14.0
hepatic fibrosis and cirrhosis (HLT), 18	1.93	8.0
increased blood pressure (HLGT), 113	1.47	16.8
hypertension and increased blood pressure NEC (HLT), 98	1.36	9.5
anaemias non-haemoltyic and marrow depression (HLGT), 203	2.42	169.8
white blood cell disorders (HLGT), 171	1.46	25.0
bullous conditions (HLT), 66	1.70	18.8
pancreatitis acute (PT), 10	1.94	4.6
lower respiratory tract inflammatory and immunologic conditions (HLT), 43	2.35	33.2
vascular disorders NEC (HLGT), 63	0.80	3.1

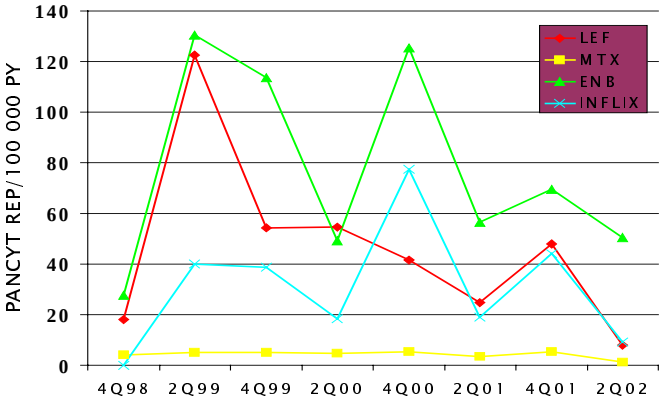
Reporting rates

Hepatic failure (FDA & IMS data)



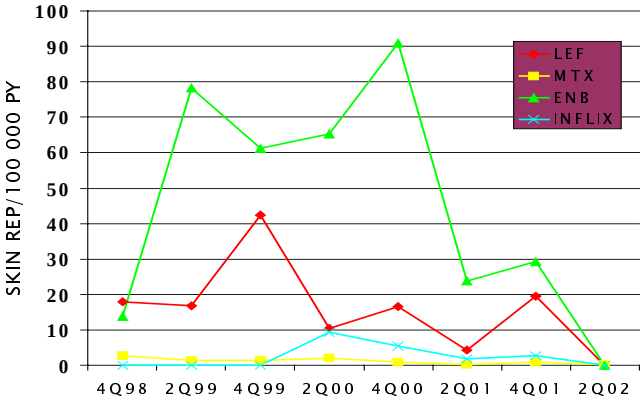
Reporting rates

Pancytopenia (FDA data, US)



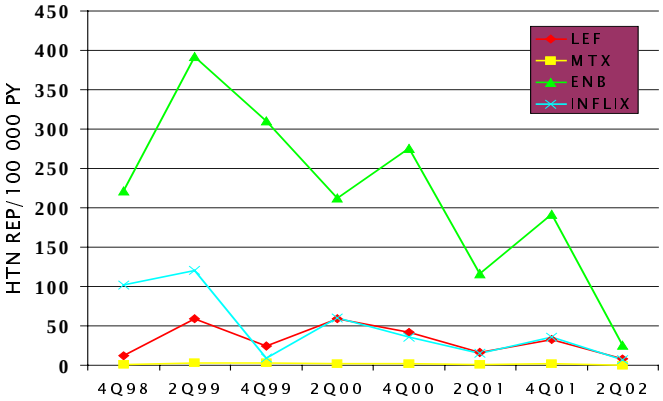
Reporting rates

Bullous conditions (FDA data, US)

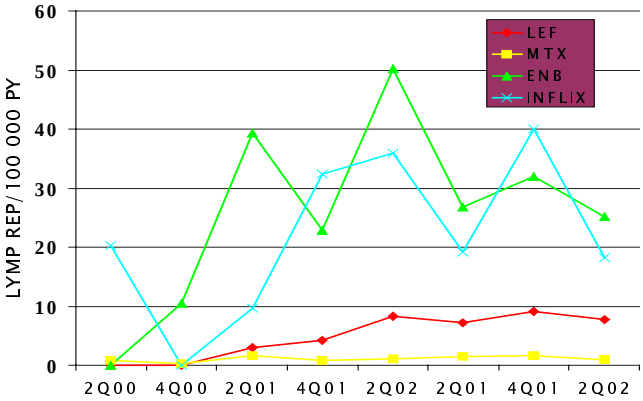


Reporting rates

Hypertension (FDA data, US)

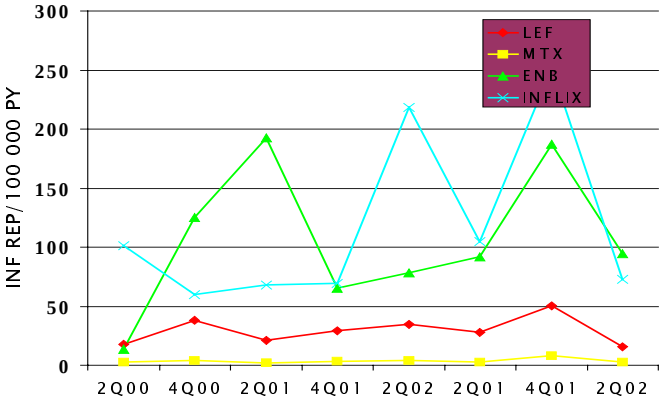


Reporting rates
Lymphoma (FDA data, US)



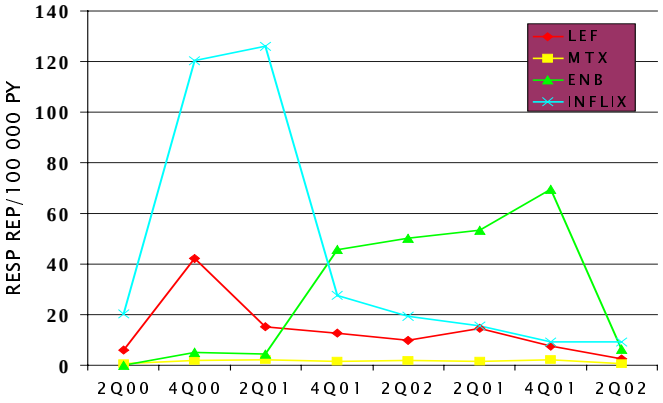
Reporting rates

Sepsis/Tuberculosis (FDA data, US)



Reporting rates

Interstitial lung disease (FDA data, US)



Reporting rates
Vasculitis (FDA data, US)

CONCLUSIONS

Part of the risk management approach to leflunomide includes the clinical safety analysis, which is performed by the Global Pharmacovigilance and Epidemiology unit. Includes in this process are the following steps:

safety signal identification

sources of safety signals (internally generated safety database, externally generated safety signals)

rapid assessment of identified signal

signal analysis

communication of findings.

All these steps were taken, and the results presented in this report represent one component of the overall process.

Based on the current PRR analysis, few signal have been generated. The PRRs for the hepatic events are less than 2.3. A PRR of 3.4 was observed for hypertensive crisis, based on 11 reports; other hypertension-related PRRs were 1.6 or less. No signals appeared in the hematologic events or severe skin events PRRs. A PRR of 4.5 was seen for necrotizing pancreatitis, although this was based on only four reports. Interstitial lung disease (ILD) was found to have a PRR of 8, based on 20 reports. A PRR of 4.9 was seen for vasculitis. Both ILD

and vasculitis have been proposed to be added to leflunomide's label. These events continue to be monitored and evaluated. No signals were observed for cancer.

Based on this analysis, few signals have been identified. Those that have been, in addition to all reports of serious events whether signal-generating or not, continue to be closely monitored using all available pharmacovigilance and epidemiologic methods.

The results of the reporting rate analysis corroborate those of the PRR analysis. No strong signals were found for leflunomide. Methotrexate was included in the analysis because it is considered the gold standard for RA therapy. Its use in reporting rate analysis, however, is questionable because of the well-known problems associated with spontaneous reporting, and a review of the graphs above demonstrate these effects quite clearly. Because methotrexate has been available for many decades (it has been used as RA therapy since 1951[7]), physicians are very familiar with its effectiveness and toxicity, and in all likelihood would not report any reactions stemming from its use. This is especially the case with rheumatologists, who are adept at using potentially very toxic drugs in daily practice.

Despite the lack of strong signals from either of these analyses, the seriousness of the reports led to more thorough epidemiologic investigations (which confirmed that the occurrence of these effects is no more frequent amongst leflunomide users than they are amongst users of other DMARDs, including methotrexate).

---

[7] Gubner R. Therapeutic suppression of tissue reactivation. I. Comparison of effects of cortisone and aminopterin. *Am J Med Sci* 1951;221:169-75.

META-ANALYSIS of LEFLUNOMIDE
Global Epidemiology

OBJECTIVE

The objective of this study was to compare the rates of adverse events seen in phase III clinical trials; specifically, leflunomide was compared to methotrexate and to sulfasalazine.

METHODS

Adverse event rates were cumulated from clinical trials US301 (placebo-controlled trial of leflunomide versus methotrexate), MN301/303/305 (placebo-controlled trial and extensions of leflunomide versus sulfasalazine), and MN302/304 (leflunomide versus methotrexate). They are presented simply as total events divided by the number persons in the different arms in the trials. The rates are presented on a L'Abbé scatter plot (line-of-identity graph) for ease and sensibility of interpretation. Points falling to the left of the line refer to rates that are higher for a comparison drug, ie, methotrexate or sulfasalazine. Points that fall to the right imply that leflunomide patients had higher rates. In addition to individual rates, adverse events that were considered 'Serious' are combined into that category, and a second category, 'Serious and Related', refers to events that were considered serious and related to the treatment drug by the treating physicians.

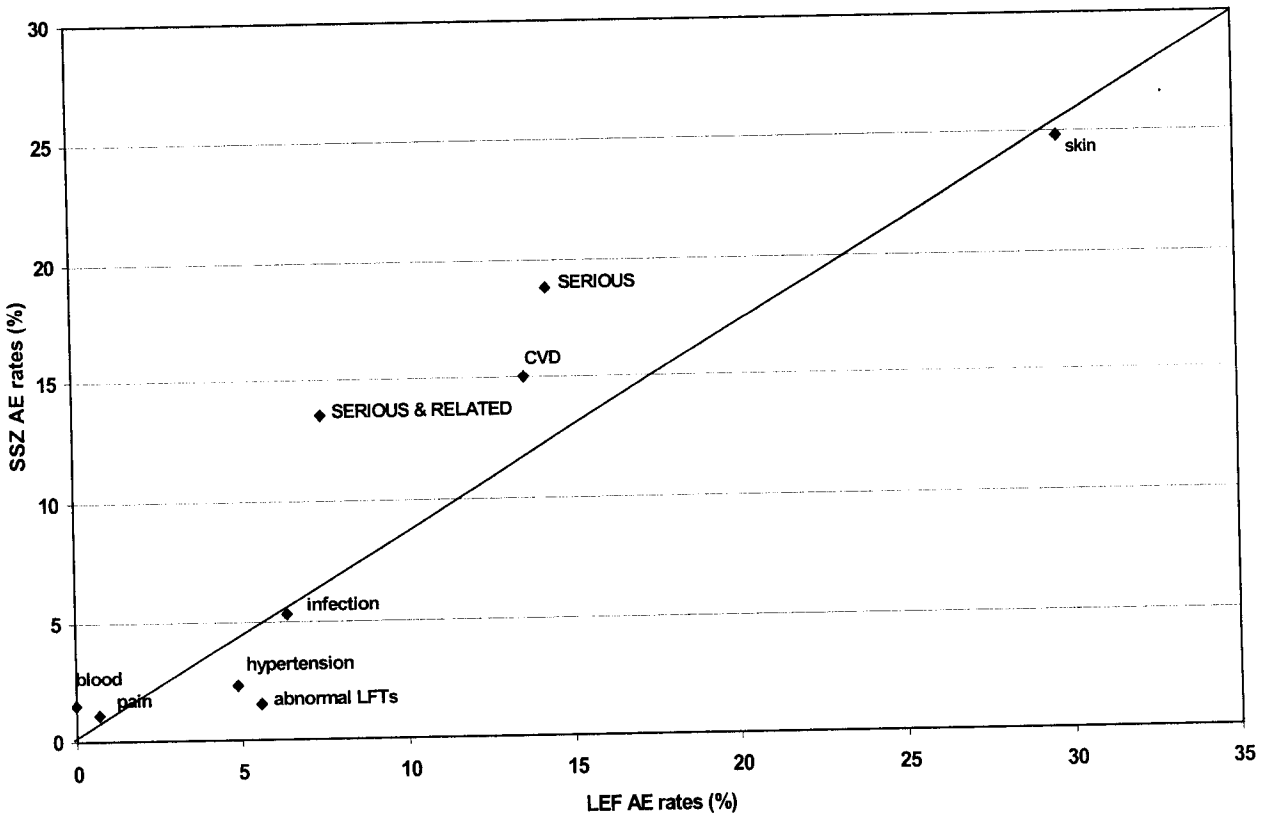
RESULTS

In general, the reported rates of AEs in the trials were quite similar, although 'Serious' and 'Serious and Related' adverse events occur more often amongst the methotrexate and sulfasalazine users. Methotrexate and sulfasalazine also had higher rates of pain, blood, and cardiovascular AEs. Skin (rash) and hypertension occurred more often amongst leflunomide users. Leflunomide had higher rates of infection and abnormal liver tests compared to sulfasalazine, and lower rates compared to methotrexate.

CONCLUSION

Using L'Abbé scatter plots to assess the rates of AEs reported in clinical trials of leflunomide, the two comparator agents (methotrexate and sulfasalazine) had comparable rates of 'Serious' and 'Serious and Related' events, as well as comparable rates of cardiovascular, blood, hypertension, abnormal liver tests, and pain AEs.

RATES OF SELECTED ADVERSE EVENTS, LEF v SSZ
data: US301, MN301/303/305, MN302/304 trials



RATES OF SELECTED ADVERSE EVENTS, LEF v MTX
data: US301, MN301/303/305, MN302/304 trials