DEPARTMENT OF HEALTH & HUMAN SERVICES

Food and Drug Administration
Center for Drug Evaluation and Research
Office of Pharmacoepidemiology and Statistical Science
Office of Drug Safety

To: Lee Simon, MD; Director
Division of Anti-inflammatory, Analgesic, and Ophthalmic Drug Products
HFD-550

From: Victor F. C. Raczkowski, MD, MS; Director
Office of Drug Safety; HFD-400

Date: January 31, 2003

Subject: NDA 20-905; Postmarketing Safety Memo
Drug: Arava® Tablets (leflunomide), 10 mg, 20 mg, 100 mg
Sponsor: Aventis Pharmaceuticals Inc.
Adverse drug experience: Hepatoxicity

Confidential: Contains IMS data; not to be used outside of FDA without clearance from IMS

Overview

Leflunomide (Arava®) is a drug used to treat rheumatoid arthritis. To date, the Office of Drug Safety (ODS) has written three reviews summarizing post-marketing reports of hepatotoxicity associated with leflunomide as captured in reports submitted to the Food and Drug Administration (FDA) through the Adverse Event Reporting System (AERS). These reviews were dated 7 March 2001, 7 November 2002, and 5 December 2002. In this memorandum I focus primarily on the review dated 7 November 2002 because over my signature I indicated that such a memorandum from me would be forthcoming.

The November 7th review focused on domestic (U.S.) reports of serious hepatic toxicity associated with the use of leflunomide. In that review, the ODS reviewers conclude that the risks of leflunomide greatly exceed its benefits.[1] The authors recommend that leflunomide be removed from the market.[2] Upon review of the data, analyses, and arguments presented in the review, I did not reach the same conclusion. Therefore, I do not concur with the recommendation.

Nonetheless, and although many of these AERS reports of leflunomide hepatotoxicity are confounded, inconclusive, or incomplete, I conclude that the totality of the data provide credible evidence that leflunomide may cause serious hepatic events, possibly mostly in patients with

[1] Memorandum from Renan A. Bonnel, PharmD., MPH., David Graham, MD, MPH, Julie Beitz, MD, and Victor Raczkowski MD, MSc addressed to Lee Simon, MD, Director of the Division of Anti-Inflammatory, Analgesic, and Ophthalmic Drug Products (HFD-550), page 2.
[2] Ibid; page 30.

underlying liver disease (e.g., alcoholic liver disease) or in patients receiving other hepatotoxic drugs (e.g., methotrexate).

The November 7th ODS review does not adequately consider the substantial medical burden imposed on patients by rheumatoid arthritis: a chronic disease that can result in joint destruction, deformity, disability, and even premature death.[3] Rheumatoid arthritis is a painful, progressive disease without a cure. Moreover, the review does not adequately consider the considerable toxicities of other treatments for rheumatoid arthritis,[4] or that management of rheumatoid arthritis often requires that a range of treatment options be available. For example, treatment options are desirable because a patient may not respond adequately to, may not tolerate the toxicities of, and/or may not comply with other treatments.[5] Adequate consideration of such factors is necessary to make an appropriate, evidence-based, regulatory action that is in the best interest of the public health.

In short, an appropriate regulatory action should consider not only the risks of leflunomide but should also consider the risks of, and suffering caused by, rheumatoid arthritis (some of which can be alleviated by leflunomide). In addition, an appropriate regulatory action should take into account the comparative risks associated with other therapies for rheumatoid arthritis. After evaluation of such factors, as well as the risks associated with leflunomide, the full range of risk-management options for the drug should be considered.

In this memorandum, I provide a very brief background of the original new drug application (NDA) for leflunomide, summarize how hepatotoxicity was addressed by the Arthritis Advisory Committee (before drug approval) and how it is reflected in the current leflunomide labeling in the United States, and summarize the three ODS postmarketing reviews of AERS reports of cases of hepatotoxicity associated with leflunomide. I then provide my comments, which are primarily focused on the three ODS memoranda.

Background of the Original New Drug Application (NDA)

The original new drug application (NDA) for leflunomide was submitted to the Food and Drug Administration (FDA) on 10 March 1998. The application was given a priority review (i.e., a 1P designation) by the reviewing Division[6] because it felt that, if approved, leflunomide would represent a significant improvement compared to marketed products in the treatment of rheumatoid arthritis. The application was taken before the Arthritis Advisory Committee on 7 August 1998. The Committee unanimously recommended that the product be approved both for the relief of the signs and symptoms of rheumatoid arthritis as well as for the retardation of structural damage in rheumatoid arthritis. FDA approved the application on 10 September 1998. As specified in the current product labeling, Arava is indicated in adults for the treatment of active rheumatoid arthritis (RA) to reduce signs and symptoms and to retard structural damage as evidenced by X-ray erosions and joint space narrowing.[7]

---

[3] American College of Rheumatology Subcommittee on Rheumatoid Arthritis Guidelines. Guidelines for the management of rheumatoid arthritis; 2002 Update. Arthritis & Rheumatism 2002;46(2):328-346.
[4] Ibid. For example, see Table 5, pages 334-335.
[5] Ibid. For example, see Figure 1, page 329.
[6] The Division of Anti-Inflammatory, Analgesic, and Ophthalmic Drug Products
[7] Package insert as of April 2000.

Summaries of Advisory Committee Deliberations over Leflunomide-Associated Hepatoxicity and Labeling for Leflunomide-Associated Hepatotoxicity

Hepatic toxicity associated with leflunomide use was identified in review of the original NDA and was discussed at the 1998 Advisory Committee meeting. The Committee was in agreement that hepatic function should be monitored during treatment with leflunomide in a way consistent with the guidelines established by the American College of Rheumatology (ACR) for methotrexate[8] and consistent with the monitoring proposed by the sponsor. In addition, the Committee was in agreement that leflunomide should not be used in patients with clinically significant liver disease.

The *Warnings* section of the current labeling for leflunomide describes hepatotoxicity and elevations of hepatic enzymes that occurred in clinical trials with leflunomide.[9] The labeling indicates that elevations of liver enzymes, primarily alanine aminotransferase (ALT) and aspartate aminotransferase (AST), occurred in significant numbers of patients and were generally reversible. More specifically, it notes that mild transaminase elevations ($\leq$2-fold the upper limit of normal [ULN]) usually resolved while continuing treatment, and that marked transaminase elevations (>3-fold ULN) reversed with dose reduction or discontinuation of treatment. The labeling recommends that, at a minimum, levels of ALT be assessed at baseline and monitored initially at monthly intervals then, if stable, at intervals determined by the individual clinical situation. The labeling goes on to provide recommendations for dose reduction, discontinuation of the use of leflunomide, and for liver biopsy based on the severity and persistence of ALT elevation.

Citing the risks of hepatic toxicity with the use of leflunomide and the role of the liver in the activation, elimination, and recycling of leflunomide, the *Clinical Pharmacology, Hepatic Insufficiency* section of the labeling recommends that leflunomide not be used in patients with hepatic insufficiency. For similar reasons the *Warnings* section of the labeling recommends against the use of leflunomide in patients with significant hepatic impairment or evidence of infection with hepatitis B or C viruses.

The *Drug Interactions* subsection in the *Precautions* section of the labeling provides data on elevation of liver enzymes in a small sample (n=30) of patients administered leflunomide concomitantly with methotrexate. It cautions that increased side effects may occur when leflunomide is given concomitantly with hepatotoxic substances.

Finally, the *Adverse Reactions* section of the labeling notes that elevated levels of ALT and AST have been associated with the use of leflunomide.

Summary of ODS Reviews of AERS reports of Hepatotoxicity Associated with Leflunomide

Three principal reviews of the postmarketing reports of leflunomide-associated hepatotoxicity have been performed to date by the Office of Drug Safety. The first review, dated 27 March 2001, summarized domestic and foreign cases of serious hepatotoxicity through 15 March 2001 These

[8] Kremer JM, Alarcon GS, Lightfoot, Robert W. Jr. et al. Methotrexate for rheumatoid arthritis. Suggested guidelines for monitoring liver toxicity. Arthritis & Rheumatism 1994;37(3):316-28.

[9] Package insert as of April 2000.

cases were subsequently included in the two subsequent ODS reviews. The second review, dated 7 November 2002, describes domestic cases of serious hepatotoxicity through 25 August 2002. The third review, dated 5 December 2002, describes foreign cases of serious hepatotoxicity associated with leflunomide through 26 September 2002. ODS was not consulted to evaluate the results of the postmarketing study conducted by Aventis titled "Post-Marketing Cohort Study of Leflunomide and Other DMARDS, a Comparative Risk Analysis." Therefore, there is no formal ODS review of the study.

ODS Review of March 27, 2001

In a memorandum dated 27 March 2001,[10] Michael F. Johnston, R.Ph., Claudia Karwoski, Pharm D. and Julie Beitz, M.D. of the Office of Drug Safety (ODS), formerly known as the Office of Postmarketing Drug Risk Assessment (OPDRA), summarize postmarketing surveillance of AERS reports for leflunomide-associated hepatotoxicity. The memorandum includes data from the time of leflunomide's approval in the U.S. in September 1998 through 15 March 2001. It describes "nine cases of hepatic failure (six fatal) and two cases of fatal toxic hepatitis associated with leflunomide." The memorandum notes that "In all of these cases, other potential confounding factors, primarily other drugs or medical history, make the exact role of leflunomide indeterminable but possibly contributory in nature. There are a variety of other non-fatal hepatic-related events (including jaundice and hepatitis) reported with leflunomide that creates suspicion of overall hepatic toxicity." At that time, ODS recommended "that hepatic-related labeling in the current Arava label be enhanced to include jaundice, hepatitis, and to mention that liver-related fatalities have occurred with Arava therapy."[11] The nine cases described in this review represent a subset of those described in the two subsequent ODS reviews.

ODS Review of 5 December 2002: Foreign cases

In a memorandum dated 5 December 2002,[12] Lauren Lee, Pharm.D, Claudia Karwoski, Pharm.D., and Julie Beitz, M.D. of ODS summarize foreign postmarketing reports of serious hepatotoxicity associated with leflunomide. The memorandum includes foreign data from its date of introduction into foreign markets (date unknown) through 26 September 2002. A total of 13 cases of liver failure were selected for further review, including one case of liver failure received directly from the Australian Therapeutic Goods Administration (TGA). As summarized by the reviewers, all 13 cases in the series met the case definition for liver failure, including three with concurrent encephalopathy and one with hepatic coma. One patient was diagnosed with toxic hepatitis and nonspecific multi-organ failure. Nine patients died, two cases did not provide outcomes, and in the remaining two cases, recovery was uncertain at the time of the report. Liver failure was the cause of death in four cases. Four of the 13 patients had underlying hepatic conditions (e.g., chronic hepatitis (2), cirrhosis, and elevated transaminases) when receiving leflunomide. Twelve of the 13 patients received at least one concomitant medication associated with or labeled for hepatic failure, fibrosis, necrosis, coma, or fatal hepatitis.[13]

[10] Addressed to Dr. Jonca Bull, Acting Director of the Division of Anti-Inflammatory, Analgesic, and Ophthalmic Drug Products (HFD-550).
[11] All quotes obtained from page 1.
[12] Addressed to Lee Simon, MD, Director of the Division of Anti-Inflammatory, Analgesic, and Ophthalmic Drug Products (HFD-550).
[13] Pages 1-2

The authors conclude that "[s]ince leflunomide has been associated with hepatotoxicity (per labeling), the progression to liver failure is a possible risk with its use. This case series of 13 liver failures, including four liver failure deaths, provides plausible evidence of such risk. In one case, a liver biopsy concluded that the hepatic injury was drug-induced. Four other cases were confounded by underlying hepatic impairment, six lacked clinical details, and two had closer temporal relationship to other drugs that have been associated with liver failure or fatal hepatic necrosis, respectively. Despite these confounding factors, a possible association between the use of leflunomide and the development of liver failure could not be excluded in these cases,. The above data are consistent with the findings in 16 US liver failure cases that addressed the possible risk of liver failure with leflunomide use."[14]

An accompanying epidemiology assessment of Australian data by David J. Graham, M.D., M.P.H. indicates that from Australian marketing approval of leflunomide in late 1999 through May 2002, 12,708 patients received authorization under the Pharmaceutical Benefits Scheme (PBS) to receive leflunomide, of which about 7,000 are currently taking the drug.[15] The exposure to leflunomide in Australia was estimated to be 12,110 person-years. The epidemiology assessment indicates that the TGA reported 18 cases of significant liver injury (4 acute liver failure, 14 other severe liver injury), nearly all as direct reports from general practitioners. ODS reviewers classified three of the four cases of liver failure as possibly drug induced. Among the 14 other severe liver injury cases, 11 were possibly drug-induced, 2 were unlikely, and one was unevaluable. The epidemiology assessment concludes that based on the Australian postmarketing data, and as adjusted for underreporting, the risk of acute liver failure is probably at least 1000 per million person-years.

Memorandum of November 7, 2002: Domestic (U.S.) cases

In a memorandum dated 7 November 2002, Drs. Bonnel, Graham, and Beitz of the Office of Drug Safety summarize domestic (U.S.) postmarketing reports of serious hepatic toxicity associated with the use of leflunomide. The memorandum includes data from the time of leflunomide's approval in the U.S. in September 1998 through 25 August 2002.

The review concludes as follows:[16] "Since approval (September 1998) through August 25, 2002, we identified a total of 102 U.S. cases of serious hepatic events with leflunomide in FDA's AERS database, of which 48 were excluded as being unrelated to leflunomide use. We identified 54 cases of serious hepatic injury that were temporally associated with the use of leflunomide. The series included 38 reports of serious hepatic injury including hepatitis, jaundice/cholestasis, and 16 reports of acute liver failure. There were nine deaths. Eight deaths were due to liver injury and one was due to interstitial lung disease. One patient underwent liver transplantation. The typical pattern of severe liver injury with leflunomide was hepatocellular or mixed in nature. Acute liver failure was unrelated to age, dose, or duration of therapy. Two-thirds of acute liver failure cases occurred in female patients."

[14] Page 2
[15] Pages 6-7
[16] Pages 1-2

"Epidemiologic analysis found that the month-specific risk of developing ALF or other severe acute liver injury remained persistently elevated for as long as leflunomide was used. The cumulative risk for the development of ALF, expressed as number needed to harm, was estimated to range from 388 to 685 after 15 months of leflunomide use, adjusted for underreporting. For the combined outcome of ALF or other severe acute liver injury, the estimated number needed to harm ranged from 107 to 188. A sensitivity analysis for ALF yielded estimates of number needed to harm ranging from 428 to 1318."

"Review of literature on hepatotoxic risks with methotrexate, another drug commonly used to treat RA, found that the major risk was the development of fibrosis, which was usually mild in degree, with no serious adverse clinical implications."

The memorandum continues: "We reviewed the experience with existing risk-management modalities and found that the available evidence indicates these are largely ineffective. To rely upon methods that have been shown to be ineffective, or for which there is not supportive evidence showing they are effective, raises ethical concerns. In addition, the occurrence of ALF is not preventable with currently available risk management strategies."

"Examination of risks and benefits of leflunomide found an absence of documented long-term benefit based on objective indices of functional ability/disability or delayed mortality. The primary measure of efficacy from pre-approval clinical trials, "%ACR20 response," was found to be based primarily on subjective criteria that are highly correlated. The risks of leflunomide greatly exceeded its benefits. We recommend that leflunomide be withdrawn from the market."

Comments on the ODS Reviews of Leflunomide

In this memorandum I focus primarily on the review dated 7 November 2002 because over my signature I indicated that such a memorandum from me would be forthcoming. The November 7th review first describes the method by which AERS cases of serious hepatotoxicity were identified. It includes an epidemiological assessment of serious hepatic risks associated with leflunomide based on domestic reports submitted to AERS as well as based on leflunomide usage data obtained from IMS Health and leflunomide data obtained from the UnitedHealth Group (UHG) and Tennessee Medicaid/Tenn Care. The review also provides an assessment of hepatotoxicity associated with methotrexate, an evaluation of the benefits of leflunomide, and comments on the effectiveness risk-management strategies.

In my comments below, I focus on some of the critical steps involved in a risk assessment (such as risk identification and risk estimation), how these steps were conducted or described in the November 7th review, and on areas where disagreements have arisen over the data, its evaluation, or its interpretation. I then comment on the assessment in the November 7th review of the efficacy and benefit-risk balance of leflunomide, risk-management strategies, ethics, and comparative toxicities of therapies for the treatment of rheumatoid arthritis. I then provide my conclusions.

Risk Identification

Risk identification involves several activities, including describing the characteristics or qualities of the risk (e.g., "*hepatotoxicity,*" "*acute liver failure,*" *"other severe liver injury"*) as well as attribution of the risk to the drug in question or to other causes (e.g., attribution of causality for hepatotoxicity to leflunomide or to other causes). These activities as related to leflunomide, particularly the attribution of causality for hepatotoxicity, are discussed below.

Risk identification is a critical step in obtaining the numerator for any subsequent quantitative risk estimation. Therefore, the factors that influence risk identification will have a major impact on any subsequent quantitative evaluations of risk and on any conclusions that can be drawn from the data. Care should be taken to describe the methods and definitions used to select cases, to avoid bias in the selection of case series, and to acknowledge uncertainties inherent in identification of the risk.

Describing the characteristics or qualities of risk: Postmarketing drug safety surveillance plays a vital role in the detection and characterization of risks, particularly for serious risks that are uncommon, that occur after long-term drug administration, or that occur in unstudied uses or unstudied populations. However, postmarketing adverse event data can be incomplete and it may be difficult to make a diagnosis with certainty (such as *acute liver failure*) from such data. These factors underscore the need to distinguish between different levels of establishing case definitions, including distinguishing between a *surveillance case-definition* and a *clinical case-definition.*[17]

The surveillance case-definition of *acute liver failure*, for example, can be thought of as the first attempt to analyze the reports based on the available diagnostic information in the report. The clinical case-definition of *acute liver failure*, however, is likely at a different level and includes the criteria essential to make a definitive diagnosis of the condition. Although data available in AERS reports are sometimes sufficient to meet the criteria for a clinical case-definition, more commonly the data are insufficient to do so. Thus, there may be substantial discrepancies between the surveillance and clinical case-definitions for a condition such as *acute liver failure* associated with the use of leflunomide, and quantitative analyses based on such a surveillance definition may differ from those based on a clinical case-definition. For example, cases classified as *acute liver failure* based on a surveillance definition may not meet criteria for a clinical case, and misclassifications may occur because of incomplete or missing surveillance data (e.g., some cases of *acute liver failure* might be more correctly classified as cases of *other severe liver injury* if more surveillance information were available).

Regardless, evaluations based on either type of case definition (surveillance or clinical) should begin with pre-specified case definitions (as was done in the November 7th ODS review). These case definitions should be applied consistently to the cases by reviewers skilled in the evaluation of clinical data.

Attribution of causality: Postmarketing drug safety surveillance plays a vital role in identifying risks that may be causally related to use of a drug.[18] However, the strength of evidence strongly

[17] Concept adapted from P. Seligman (personal communication)

influences the certainty with which one can conclude that a medication is causally associated with a particular adverse event.[19] For adverse events that are very uncommon, observational studies may be the most feasible way to establish a causal association. In some cases, well-designed and executed observational studies (e.g., cohort studies) can be persuasive in demonstrating a causal association between a drug and an adverse event, particularly if the findings are replicated under different conditions of study.

During the evaluation of any individual case report, however, it is often difficult or impossible to conclude with certainty that a particular drug caused a particular adverse event in the patient, particularly when data are incomplete, there are confounding factors that could have resulted in the adverse event of interest, or there is a non-negligible background rate of the adverse event. Nonetheless, such individual case reports considered in aggregate may be persuasive despite the limitations of any individual report. In such circumstances, it is important to use clear criteria for causality, to openly describe the limitations of the data, and to accurately reflect differences in opinion about causality assessment.

a) Causality assessments: Issues with the classification scheme

The November 7th review clearly articulates criteria by which hepatic adverse events would be judged as *probably/likely*, *possibly*, or *unlikely* related to leflunomide use.[20] The classification, however, is incomplete and would have been enhanced by use of at least one more category such as *unclassified*. At least one such additional category would allow more accurate capture of data in those circumstances when information in the reports was missing, insufficient, contradictory, etc. In the absence of such a category or categories, inclusion of such cases in the epidemiological assessment could lead to misleading or inaccurate estimates of risk.

The November 7th review utilizes some of the standard criteria that were developed by the World Health Organization (WHO) to classify causality. However, the review utilizes an adaptation of the classification scheme that does not include important parts of the WHO classification scheme (such as WHO categories that capture cases with missing or insufficient data, see preceding paragraph). Moreover, the classification scheme is adapted in a way that has the potential to be confusing. That is, a criterion for categorizing an adverse event as *possibly* related to the use of leflunomide could be easily confused with a similar criterion used to classify an adverse event as *probably/likely* related to the use of leflunomide.[21]

Specifically, for an event to be classified as ***probably/likely*** attributable to leflunomide by the ODS reviewers, it had to meet the following criterion (among others): **"The event is <u>unlikely</u> attributed to concurrent disease or other drugs**." However, an event would be classified as

---

[18] Particularly for serious risks that are uncommon, that occur after long-term drug administration, or that occur in unstudied uses or unstudied populations.

[19] Multicenter, randomized, fully-blinded, concurrently controlled studies are generally accepted as providing the strongest evidence to support a causal relationship between a drug and an adverse event, particularly when the trials are replicated under different conditions.

[20] Pages 5-6

[21] The WHO criteria for *possible* associations do not include the requirement that confounding factors *need not be the most likely explanation for the adverse event*. Notably, the common meaning of *possibly* is similar to the *non*-adapted WHO criteria for causality assessment.

possibly related to leflunomide, if it met the following criterion (among others): "**There are other factors present that could plausibly have contributed to liver injury, but [they] were not the most likely explanation for the adverse event."** [22]

Thus, for an event to be classified as *probably/likely* related to leflunomide use by the reviewers, it needed to be <u>*unlikely*</u> *attributed to concurrent disease or other drugs*. However, for the event to be classified as *possibly* related to leflunomide use such factors (e.g., concurrent disease or other drugs) *needed not [be] the most likely explanation for the adverse event.* The distinction, if there is one, between *unlikely* and *not the most likely* in these two different categories is a subtle one, both logically and semantically. Hence, the distinctions between cases classified as *probably/likely* and *possibly likely* have the potential to be blurred as well.

Consequently, it remains unclear whether during the analysis of the cases (summarized in the November 7th review) by the multidisciplinary group of scientists,[23] the individuals voted on the basis of the reviewers' *adapted* WHO criteria or on the basis of the common meaning of the word *possibly* (i.e., *capable of happening without contradicting circumstances*). If individuals voted using the latter common meaning, then the attribution by the multidisciplinary group of scientists summarized in Table 12 of the review may be an inflated causality assessment. That is, many of the *possibly* related votes by the group should instead be classified as *unlikely* related to leflunomide. These comments reflect the importance of establishing clear criteria for causality assessments and to ensure that such criteria are fully understood by those performing such assessments. Otherwise, the validity of the causality assessment may be questionable.

b) Causality assessments: Disparities between assessments by the reviewers and the group

Despite the ambiguity in nomenclature and classification cited above, the disparity is notable between the causality assessment of the16 cases of *acute liver failure* by the authors of the review and that by the multidisciplinary group of scientists. Whereas the authors of the review classified 12 cases as *probably*-related, 4 cases as *possibly*-related, and 0 cases as *unlikely*-related to the use of leflunomide, the majority of the multidisciplinary group of scientists (which included three gastroenterologists and two voting authors of the review) classified only 3 cases as *probably*-related to the use of leflunomide, 6 cases as *possibly*-related, and 1 case as *unlikely*-related to leflunomide. The remaining 6 cases did not have a clear majority vote for any category, though the votes tended overall toward being *possibly*-related to leflunomide.

There are several possible explanations for such disparity. One possibility is that it may reflect individual differences in the weights assigned to various pieces of evidence among the reviewers. For example, it could reflect a difference in frames of reference for those accustomed to evaluating clinical trial data and those accustomed to evaluating observational or epidemiological data. Another explanation for the discrepancy could be systematic differences in the way that missing data or confounding variables are considered by different individuals.

[22] See pages 5-6 of the November 7th review.
[23] Results summarized in Table 12 on page 18.

But perhaps the most likely explanation for the disparity could be the way that the concept of *implied causality* is being applied to the evaluation of spontaneous reports of suspected adverse drug reactions (suspected ADRs). This principle is articulated in the description of standards for expedited reporting of single cases of serious, unexpected adverse drug reactions by the International Conference on Harmonisation (ICH). ICH guideline E2A states: "For purposes of reporting, adverse event reports associated with marketed drugs (spontaneous reports) usually imply causality."[24]

The Report of CIOMS Working Group V elaborates: "A basic principle upon which spontaneous reporting systems have been built and analyzed over the past decades is the assumption of at least a "possible" causal relationship between the event(s) reported and one or more specified drug products (i.e., it is a suspected ADR). In other words, the voluntary nature of the initial communication reflects an index of suspicion on the part of the reporter regarding the role of one or more products."[25]

Apparently, a similar concept was applied by the ODS reviewers in their assessment of causality. A statement in the epidemiological assessment indicates that: "The physicians who submitted these case reports believed that leflunomide was causally responsible or contributory."[26]

However, systematic application of such reasoning to the evaluation of individual case reports, particularly in the presence of confounding factors, missing data, or contradictory data, will lead toward a systematic tendency to attribute a causal relationship of an adverse event to the drug when none may in fact exist. As mentioned above, it is often extremely difficult or impossible to conclude with certainty that a particular drug caused a particular adverse event in an individual case report. Thus, the routine application of such reasoning to the evaluation of individual case reports may overstate the degree of any causal relationship.

Moreover, MedWatch form (FDA form 3500), a form commonly used to submit spontaneous reports of serious adverse events to FDA, prominently carries a statement that undercuts such reasoning: "Submission of a report does not constitute an admission that medical personnel or the product caused or contributed to the event." The statement could be interpreted by reporters to submit the form even if they do not suspect a causal relationship between the drug and the adverse event. Thus, the assumption that "The physicians who submitted these case reports believed that leflunomide was causally responsible or contributory" may not be entirely true.

But since the goal is to establish a true (i.e., non-biased) risk assessment, cases should be evaluated on the strength of their evidence with clear and explicit acknowledgement of uncertainties, confounding factors, contradictory data, or missing data. To illustrate, the ODS review of 5 December 2002 explicitly states in the Executive Summary: "Four...cases [of

[24] ICH E2A: Clinical Safety Data Management: Definitions and Standards for Expedited Reporting.
[25] Current Challenges in Pharmacovigilance: Pragmatic Approaches. Report of CIOMS Working Group V (2001). Council for International Organizations of Medical Sciences, Geneva.
[26] Page 20

leflunomide-associated hepatotoxicity] were confounded by underlying hepatic impairment, six lacked clinical details, and two had closer temporal relationship to other drugs that have been associated with liver failure or fatal hepatic necrosis, respectively....Twelve of the 13 patients received at least one concomitant medication associated with or labeled for hepatic failure, fibrosis, necrosis, coma, or fatal hepatitis."[27]

Similarly, the ODS review of 27 March 2001 explicitly states in the Executive Summary: "In all of ...[the eleven cases of leflunomide-associated hepatotoxicity]...other potential confounding factors, primarily other drugs or medical history, make the exact role of leflunomide indeterminable but possibly contributory in nature."

The ODS review of 7 November 2002, however, has no similar explicit statement in the Executive Summary. Moreover, the epidemiological assessment explicitly acknowledges the intentional introduction of some degree of systematic bias into the review: "The reason that most 'possible' reports were classified as such is because clinical and laboratory information provided in the reports was incomplete. To exclude these reports would introduce a bias favorable to the drug and its manufacturer because were the missing information supplied, some proportion of these reports would be upgraded to the 'probable' category. Also, excluding these reports presumes that leflunomide had no role in causation (which has not been shown) and would have the effect of rewarding poor quality reports."[28] Thus, the epidemiological assessment takes the explicit position that scientific accuracy should be sacrificed as a method by which to exert leverage over manufacturers.

However, an evidence-based approach of how to deal with such incomplete information would be to identify such cases clearly, and then to perform a sensitivity analysis with those cases included in the various causality categories (e.g., to perform both a sensitivity analysis that includes "best-case" and "worst-case" evaluations).

Finally, the assertion that "some proportion of these reports would be upgraded to the 'probable' category" appears to be without foundation. It is conceivable that additional information would in fact downgrade some of such reports to the 'unlikely' category.

In summary, the goal of causality evaluation is to establish a true (i.e., non-biased) risk assessment. It is therefore essential that such causality evaluations be conducted using clearly established rules for handling data (e.g., missing data) that are applied consistently by reviewers skilled in the evaluation of clinical data. Bias should not be introduced intentionally into such evaluations.

c) Causality assessments: Incomplete sensitivity analysis[29]

The sensitivity analysis of the vote of the multidisciplinary group of scientists as reflected in the epidemiological assessment is incomplete. Because it is incomplete, the analysis conceals the fact that the majority of the multidisciplinary group of scientists classified only three cases

[27] Pages 1-2
[28] Page 20
[29] Pages 18-19.

as *probably* related to leflunomide, in contrast to the reviewers that classified 12 cases as *probably* related.

Moreover, the analysis appears to have been performed in a way that differs from the analyses reflected in the tables based solely on the reviewers' causality assessment. For example, Tables 6, 7, 8, 9, 10 and 11 provide estimates of the "number needed to harm" (NNH) under various scenarios. All of these tables are based solely on the reviewers' causality assessment and all explicitly include a calculation of NNH based only on the *probably*-related cases.

In contrast, the epidemiological assessment appears to neglect a major conclusion reached by the multidisciplinary group of scientists. There are no tables that capture the majority vote by the multidisciplinary group of scientists for *probably*-related cases as there were for the reviewers' assessment. If such tables were constructed, the "numbers needed to harm" calculations based on the group's *probably*-related assessment would have been markedly higher than those estimated based on the reviewers' assessment.

Risk Estimation

Risk estimation includes activities such as quantifying the degree of risk. Risk may be estimated with a variety of measures, including the *crude rate*, *cumulative rate*, and *hazard rate*.[30] The *crude rate* is the number of patients experiencing a certain adverse event divided by the number of subjects of patients exposed to the drug, regardless of duration of use. The *cumulative rate* is the probability that an adverse event of a specified type will occur by a particular point in time. The *hazard rate* is the probability of an adverse event during a specified interval of time, assuming that the patient has been exposed and followed to the beginning of the interval.

The epidemiological assessment in the review makes attempts to estimate reporting rates, quantify risk, assess hazard rates, and calculate cumulative risks for leflunomide-associated hepatotoxicity. Also, some of the assumptions of the epidemiological assessment are stated in the text (e.g., in the assessment of hazard rates, cases of unknown duration of exposure to leflunomide are distributed in proportion to those with known duration[31]). However, the epidemiological assessment lacks a summary of how such assumptions affect the interpretation of the results, particularly if the assumptions are incorrect or are violated.[32] For example, the available data on underreporting in AERS may be insufficiently precise and insufficiently generalizeable to all drugs and situations to allow precise quantitative adjustments for underreporting. Thus, while the sensitivity analysis using different reporting efficiencies is a desirable feature of the epidemiological analysis,[33] these analyses may yield misleading, invalid, or imprecise figures in the calculation of the "numbers needed to harm" as reflected in Tables 10 and 11.[34]

[30] O'Neill RT (1988). Assessment of Safety. In: *Biopharmaceutical Statistics for Drug Development.* Peace KE (Editor), pp 543-604. Marcel Dekker, Inc, New York.
[31] Page 14.
[32] Such discussions of "limitations" of a study or of analyses are common in peer-reviewed, evidence-based biomedical journals.
[33] Pages 16-17.
[34] Page 16.

For example, over 1/3 of the cases of *other severe liver injury* were of unknown duration (28 known duration;10 unknown).[35] Because these cases represent such a large percentage of the data, the validity of the conclusions reached by the epidemiological assessment may be undermined if the assumptions about distributing the cases are incorrect or are violated. For example, if the hazard of serious hepatic injury associated with leflunomide toxicity is in fact greatest upon initiation of therapy and drops significantly afterwards (to zero, in the most extreme case), then the "distribution...[of these cases] ...in proportion to those with known duration"[36] as was done in the epidemiological assessment may lead to misleading assessments of risk. In such a case, the risk estimation would be particularly misleading if the purported "known" cases of long duration of exposure were in fact cases that were not attributable to leflunomide but were cases of hepatotoxicity due to other causes (e.g., "false positive" cases). However, the epidemiological assessment did not systematically or explicitly address such issues.

Similarly, the figure of month-specific hazard rates of reported acute liver failure (Figure 4) and the interpretation of these data offered other opportunities to discuss limitations of methodology (or of consequences when assumptions may be incorrect).[37] In the epidemiological assessment, the following unqualified statement is made: "Of note, the lower bound of the 95% confidence interval for ALF [acute liver failure] excluded 1 per million person-years, which is the estimated background rate of ALF in the general population."[38] However, this conclusion is based on surveillance case-definitions (see above) and different conclusions might be reached if the analysis had been based on clinical case-definitions.

Moreover, and perhaps of greater significance, this analysis was based on the reviewers' causality assessment (12 *probably/likely*-related, 4 *possibly*-related) which was acknowledged in the epidemiological assessment to be an intentionally inflated causality assessment. This conclusion would likely differ if the cases with missing or incomplete data were handled differently in the analysis. And the conclusion would certainly differ if the analysis was based on the majority vote of the multidisciplinary group of scientists (who classified only 3 cases as *probably*-related to the use of leflunomide).

Thus, the figure that shows a relatively constant month-specific hazard rate of acute liver failure over 15 months (figure 4) may be very misleading if the majority vote approximates the true state. For example, the hazard analysis would not have been able to extend beyond 6 months because all three of the *probable* cases (by the majority vote) occurred within the first six months of leflunomide treatment (cases #2, #11, and #16).

In conclusion, the estimated rates for *acute liver failure* or for *other severe liver injury* are known with substantially less accuracy and precision than reflected in the text or as summarized in the executive summary. The reader must scrutinize the body of the review to realize that the epidemiological assessment yielded values of the number needed to harm (NNH) that ranged over at least one order of magnitude. Specifically, Table 10 shows that based on the epidemiological assessment, the cumulative risk for the development of *acute liver failure* (ALF), expressed as the

[35] Page 14.
[36] Page 14.
[37] Pages 14-16.
[38] Pages 15-16.

number needed to harm, was estimated to range from 194 to 2408 after 15 months of leflunomide use, adjusted for underreporting.[39] Moreover, this range of 194 to 2408 would be substantially broader (e.g., higher numbers of patients needed to harm) if the epidemiological analysis had incorporated a systematic evaluation of the limitations of the methodology (e.g., if the limitations of the analyses had been systematically evaluated, if these cases represent bona fide cases of acute liver failure, etc.). This range contrasts with the remarkably precise values of 388 to 685 presented in the executive summary.

Similarly, Table 11 shows that based on the epidemiological assessment, the cumulative risk for the development of the combined outcome of *acute liver failure* and *other severe acute liver injury*, expressed as the number needed to harm, was estimated to range from 54 to 702 after 15 months of leflunomide use, adjusted for underreporting. As above, this range of 54 to 702 would be substantially broader (e.g., higher numbers of patients needed to harm) if the epidemiological analysis had incorporated a systematic evaluation of the limitations of the methodology (e.g., if the limitations of the analyses had been systematically evaluated). This contrasts with the remarkably precise values of 107 to 188 presented in the executive summary.

Thus, the limitations of the data and analyses were not fully explored or articulated in the epidemiological analysis or in the review. Consequently, it is likely that the estimates provided in the executive summary represent the most-extreme (e.g., worst-case) estimates of rates and represent rates that that are overly precise given the limitations and assumptions upon which the analyses were performed.

Efficacy and Benefit-Risk Balance of Leflunomide

The November 7th review makes comments on the efficacy and benefit-risk balance of leflunomide.[40] Several arguments are made in support of the authors' view that the benefit-risk profile of leflunomide is unfavorable. However, these arguments are unpersuasive based on the data provided in the review.

For example, the November 7th review asserts that "This lack of persistent use is important to recognize because it suggests that leflunomide is not well-tolerated (side effects), or else isn't very effective."[41] While such an explanation is a possibility, use of such persistence data to argue that a drug such as leflunomide is ineffective is an extrapolation that goes well beyond the data. Other major factors (e.g., cost, continued availability of leflunomide in health plans or in formularies, convenience, physician familiarity, extent of advertising) were not considered or mentioned in the review as possible explanations. Moreover, no comparable data (e.g., no "control" data) were shown for other newly marketed drugs to form any basis for comparison.

The November 7th review provides data on the persistence of use of methotrexate,[42] but comparisons between methotrexate and leflunomide are of unclear relevance and may even be invalid given differences between the two drugs. For example, methotrexate been marketed for

[39] Page 16.
[40] Pages 23-25.
[41] Page 23.
[42] Table 19, page 23.

many more years than leflunomide, and any comparison between the two is likely to be severely confounded by secular (time-related) trends in drug usage patterns for reasons unrelated to either the safety or efficacy of either drug (e.g., cost, continued availability of leflunomide in health plans or in formularies, convenience, physician familiarity). In addition, another important confounding factor is that methotrexate has other uses besides its use in rheumatoid arthritis. Comparisons of drugs (like methotrexate), which have many uses and that are relatively well understood, to newer drugs (like leflunomide), with more limited uses are of uncertain relevance and may be confounded. Thus, comparisons of the efficacy of methotrexate with that of leflunomide based on persistence of use is an extrapolation that goes well beyond the data presented in the November 7th review.

Leflunomide labeling states in the Clinical Studies section: "No consistent differences were demonstrated between leflunomide and methotrexate or between leflunomide and sulfasalazine." This conclusion was reached by FDA from its intensive, multidisciplinary reviews of the design, conduct, and results of randomized, controlled clinical trials. The results of these trials were also evaluated by experts of the Arthritis Advisory Committee in a public advisory committee meeting, and these experts reached a similar conclusion. Thus, clinical trial data have not shown that methotrexate is superior to leflunomide in terms of efficacy**.** Such randomized, controlled trials are a much higher stronger level of evidence than "persistence of use" data, particularly if the use data are not appropriately controlled or if confounding factors are not taken into account (e.g., adjusted for) in some fashion in the analysis.

The November 7th review also asserts that "....virtually all the endpoint measures upon which efficacy was based were subjective in nature and represent surrogates for what might be considered true benefit, such as delay/prevention of permanent loss of significant physical functioning, delay/prevention of serious disability, and delay in premature mortality."[43] As described below, this statement appears to confuse the terms "subjective endpoint" with that of a "surrogate endpoint."[44]

Subjective endpoints (outcome measures): The statement cited in the preceding paragraph seems to imply that because pain, suffering or other symptoms of patients with rheumatoid arthritis may be subjective, alleviation of such symptoms is not of important clinical significance. However, FDA often approves drugs based on symptomatic and/or subjective improvement. Moreover, evaluation of subjective endpoints can be valid and reliable, particularly when evaluated in randomized, blinded (masked), clinical trials.

Discussions of the importance of alleviation of pain and suffering with patients and/or their healthcare providers indicate that alleviation of pain and suffering are of enormous clinical importance and are a worthy goal of drug therapy. In some cases, patients may even choose alleviation of pain and suffering over prolongation of survival or improvement of long-term disability. Thus, Figure 7 of the November 7th review provides a perspective on the benefit-risk

[43] Page 24.

[44] A *subjective endpoint* is an endpoint (outcome measure), such as pain, that is only perceived by the patient and that is not perceptible to the senses of another person. A subjective endpoint contrasts with an *objective endpoint* that is perceptible to the external senses of another person.

balance of leflunomide that understates the importance of alleviating subjective outcomes such as pain and suffering.

Improvement of pain or suffering can also lead to improved compliance with taking medication. Conversely, a regimen to take medication that improves long-term disability or survival but that does not have an impact on pain or suffering may have low patient compliance. Therefore, the long-term benefits of the drug (improvement in survival or disability) may never be realized in patients who are not compliant with taking the medication.[45]

In conclusion, making (or keeping) drugs available that alleviate pain and suffering is an important part of CDER's goal of improving public health, and the clinical importance of such alleviation should not be minimized or disregarded.

Surrogate endpoints (outcome measures): This statement cited above in the November 7th review also seems to imply that endpoints as used to assess the status or progression of rheumatoid arthritis (e.g., radiographically identified structural damage to joints) are also not of important clinical significance. A full discussion of surrogate endpoints is beyond the scope of this memorandum. However, in some cases (e.g., treatment of hypertension or glaucoma) FDA accepts surrogate endpoints (e.g., reduction of blood pressure, reduction of intraocular pressure) as sufficient evidence for marketing approval. In other cases, FDA can approve drugs intended to treat serious and life-threatening diseases based on an effect on a surrogate endpoint, provided that certain criteria are met and that there is a commitment to define the actual clinical benefit of the agent in studies completed after marketing.[46] Moreover, the American College of Rheumatology, through its committees, has defined and standardized the diagnosis of rheumatoid arthritis and various assessments of disease activity. These have been widely and successfully used, so that clinical experts, clinical investigators, and FDA use them to evaluate drugs used to treat rheumatoid arthritis.

Use of endpoints for rheumatoid arthritis such as "prevention of structural damage" have been discussed publicly at meetings of the Arthritis Advisory Committee and have been incorporated into an FDA guidance: "Guidance for Industry; Clinical Development Programs for Drugs, Devices, and Biological Products for the Treatment of Rheumatoid Arthritis (RA);" February 1999.[47] This guidance represents the collective views of three FDA centers that evaluate products for rheumatoid arthritis: the Center for Drug Evaluation and Research (CDER), the Center for Biologics Evaluation and Research (CBER), and the Center for Devices and Radiological Health (CDRH). Moreover, the guidance notes that the claim of prevention of structural damage would be submitted for an agent that has been shown (previously or concomitantly) to be effective for one of the other claims (e.g., prevention of disability).[48] Thus, the clinical importance of "prevention of structural damage" should not be minimized or disregarded, particularly when the findings are consistent with other outcome measures of clinical improvement.

---

[45] Examples include antihypertensive drugs and cholesterol-lowering agents, which because of their side effects and because they do not alleviate pain or suffering, can reduce a patients motivation to comply with treatment.
[46] See 21 CFR 314, subpart H and 21 CFR 601, subpart E.
[47] Available at www.fda.gov/cder/guidance/index.htm under "Clinical/Medical"
[48] The guidance also discusses endpoints such as "major clinical response," "complete clinical response," "remission," and "prevention of disability."

Risk Management:

Based on evaluations of some drugs, the November 7th review cites data and takes the position that use of risk management tools (including labeling changes, educational efforts, and restricted distribution programs) short of drug withdrawal are largely ineffective.[49] However, it's not clear the extent to which these results are generalizeable to other drugs or other situations.

Publicity and re-labeling, for example, can have a significant impact on physician and patient behavior. For example, a study of the co-administration of terfenadine with either ketoconazole or erythromycin (known in combination to prolong the QT interval and to be a risk factor for torsades de pointes) showed that there were rapid and significant declines in the rates of filling prescriptions for either erythromycin or ketoconazole within 2 days of prescriptions for terfenadine in three paid pharmacy claims databases.[50] These declines coincided with publicity concerning the safe use of terfenadine. Although the results of this study did not show perfect "compliance" with FDA recommended labeling (i.e., a coadministration rate of zero), the rate of concomitant use of ketoconazole and erythromycin with terfenadine fell by 80% or more. Other authors have published complementary results.[51, 52] Thus, this example of terfenadine is a case where large, but not perfect, results were obtained with publicity and labeling. Had terfenadine not been withdrawn from the market, further improvements in these rates would have been desirable. Therefore, the goal in such cases of partial success could be framed not as a failure, but as an opportunity to build further on these successes of labeling and publicity to further optimize "compliance" with recommended labeling.

The use of spironolactone provides another example of a positive impact of physician education on drug use.[53] The Randomized Aldactone Evaluation Study (RALES) showed a reduction in mortality in patients with severe heart failure randomized to receive spironolactone, a long-marketed drug with multiple manufacturers and generic versions. The authors concluded that prescribing of the single ingredient spironolactone for cardiovascular diseases increased markedly after publication of the RALES Trial. Drug company and other promotion of spironolactone for this clinical benefit was negligible, suggesting that physician prescribing can be materially influenced by significant publications in the medical literature.

Cisapride (a drug withdrawn from the U.S. market because of risks of QT prolongation and torsades de pointes) is cited in the November 7th review as a case where, "despite labeling efforts including bolded, boxed warnings, labeled contraindications, and 'Dear HCP [health-care professional]' letters, there was no change in the unsafe use of the drug."[54] However, another study found that "A highly publicized letter sent in June 1998 was associated with a notable decline (58%) in the concomitant dispensing rates of with explicitly contraindicated drugs but not

[49] Pages 25-27.
[50] Burkhart GA, Sevka MJ, Temple R, Honig PK. Temporal decline in filling prescriptions for terfenadine closely in time with those for either ketoconazole or erythromycin. Clin Pharmacol Ther 1997;61:93-6
[51] Carison AM, Morris, LS. Coprescription of terfenadine and erythromycin or ketoconazole: An assessment of potential harm. Journal of the American Pharmaceutical Association, 1996;Vol NS36, No.4, 263-269.
[52] Thompson D, Oster G; Use of terfenadine and contraindicated drugs. JAMA 1996;275:1339-1341.
[53] Karwoski CB, Trontell A, Beitz J, Swann J. Increased prescriptions for spironolactone following publication of the RALES trial. Submitted for publication.
[54] Page 25.

in the concomitant dispensing of cisapride with the example or implied drugs. An earlier letter, which had been explicit but was accompanied by less publicity, had no measurable effect on this study's measure of coprescription, nor did a later letter that emphasized comorbidities.[55] The authors concluded that explicit, well-publicized drug warnings can change physician behavior. Thus, even in the case of cisapride cited in the November 7th review as a failure of risk-management efforts, other studies have found that risk-management efforts can indeed have impacts on behavior.

Thus, depending on the ultimate benefit-risk assessment of leflunomide, a whole range of risk/management options could be considered. These include labeling changes (e.g., black box warnings, medication guides), educational efforts, restricted use or distribution (e.g., prescription only to adequately informed patients, prescribing only to patients with particularly aggressive rheumatoid arthritis, prescription only by physicians knowledgeable by training or experience in the treatment of rheumatoid arthritis and in the benefits and risks of leflunomide). Withdrawal from the market could be considered if the risk-benefit profile is ultimately deemed to be unacceptable.

If risk-management options (short of withdrawal) are pursued, it would be highly desirable that the goals of the program be clearly articulated. For example, demonstration of an actual change in prescribing behavior could be one goal. However, alternatives to such a behavior-based goal could be to ensure that patients are adequately informed before and during leflunomide therapy, and that physicians have the best information available to make an appropriate therapeutic decision in consultation with their patients (e.g., to facilitate educated and informed decision making on the part of patients and physicians). In other words, the goal need not be for FDA to use its regulatory authority *to control* the patient-physician interaction, but rather *to facilitate* an appropriate and knowledgeable interaction.

Thus, if risk-management options (short of withdrawal) are pursued, it would be highly desirable to evaluate the effectiveness of the individual components of the risk-management program as well as the effectiveness of the overall program. For example, it's possible that when assessed individually, specific risk-management "tools" are found to be relatively ineffective. However, when the overall program is assessed, the program has the desired (or improved) outcome. The publication cited above on cisapride, for example, suggests that additive or synergistic effects may be important in changing behavior by health care providers, even when tools in isolation are relatively ineffective. If after such an evaluation the risk-management program is partially successful, subsequent iterations of the risk-management program could focus on improving areas of weakness or of sub-optimal results.

Ethics:

The November 7th ODS review raises purported "ethical considerations" raised by the continued marketing of leflunomide.[56] However, this section assumes that the November 7th review has made a convincing case for an unfavorable benefit-risk profile of leflunomide and a convincing

[55] Weatherby LB, Nordstrom BL, Fife D, and Walker AM. The impact of wording in "Dear doctor" letters and in black box labels. Clin Pharmacol Ther 2002;72:735-42.
[56] Pages 27-29.

case that risk management interventions short of drug withdrawal would be futile. However, given the issues discussed above, both of these assumptions are questionable. Stated differently, employees in all branches of the Center for Drug Evaluation and Research (CDER) share similar ethical goals of respect for persons, beneficence, and justice. These ethical goals are not the relevant issue. Rather the relevant issues that need resolution are related to different perspectives on the benefits-risk profile of leflunomide as well as to different perspectives of the potential effectiveness of risk-management interventions for the drug.

I concur with the memorandum of 20 December 2002 by Dr. Paul Seligman who states that "Efforts to minimize the risks of therapy and to understand problems that may arise post-marketing are not 'experiments,' as indicated in the memo, but are ethical and humane approaches to evaluate and ensure the best and most appropriate use of drugs."

Comparative Toxicities of Therapies for the Treatment of Rheumatoid Arthritis

In terms of comparative treatments for rheumatoid arthritis, the November 7th review focuses on hepatotoxicity associated with methotrexate. However, an appropriate regulatory action should take into account the comparative toxicities associated with other therapies for rheumatoid arthritis, and how those therapies are used**.** The November 7th review does not adequately consider the considerable toxicities of other treatments for rheumatoid arthritis,[57] or that management of rheumatoid arthritis often requires that a range of treatment options be available. For example, treatment options are desirable because a patient may not respond adequately to, may not tolerate the toxicities of, and/or may not comply with other treatments.[58] Adequate consideration of such factors is necessary to make an appropriate, evidence-based, regulatory action that is in the best interest of the public health.

While a comparison of hepatic toxicity of methotrexate with that of leflunomide is of interest (as was performed in the November 7th review) , the larger question is how the **total** serious toxicities of other therapies compare with the **total** serious toxicities of leflunomide and other drugs used to treat rheumatoid arthritis. For example, serious toxicities associated with methotrexate include myelosuppression, hepatic fibrosis, cirrhosis, pulmonary infiltrates, and pulmonary fibrosis. Serious toxicities associated with salicylates and nonsteroidal antiinflammatory drugs include gastrointestinal ulceration, gastrointestinal bleeding, and hepatic toxicity. Sulfasalazine is associated with serious myelosuppression.. Glucocorticoids are associated with serious cases of hypertension, hyperglycemia, and osteoporosis. A serious toxicity of hydroxychloroquine is macular damage. The toxicities of these drugs as well as of other agents used to treat rheumatoid arthritis have been summarized by the American College of Rheumatology.[59] Whether the toxicities of leflunomide exceed the toxicities of these other agents remains an open question.

In conclusion, an appropriate regulatory action for leflunomide should take into account the comparative risks associated with other therapies for rheumatoid arthritis, including the total

[57] Ibid. For example, see Table 5, pages 334-335.
[58] Ibid. For example, see Figure 1, page 329.
[59] American College of Rheumatology Subcommittee on Rheumatoid Arthritis Guidelines. Guidelines for the management of rheumatoid arthritis; 2002 Update. Arthritis & Rheumatism 2002;46(2):328-346

serious toxicities of each agent as well as the profile of less-serious toxicities that could affect patient compliance with therapy.

Conclusions:

Domestic and foreign spontaneous adverse event reports submitted to AERS provide credible aggregate evidence that leflunomide may cause serious hepatic events. Such a finding is not entirely unexpected given that at the time of approval, leflunomide was associated with hepatotoxic adverse events (that were typically less serious). In individual cases, it is extremely difficult or impossible to draw definitive evidence of causality because many of the cases are confounded, inconclusive, or incomplete. However, the totality of the data appear to indicate that serious hepatotoxic adverse events occur with the use of leflunomide, possibly more so in patients with underlying liver disease (e.g., alcoholic liver disease) or in patients receiving other concomitant hepatotoxic drugs (e.g., methotrexate).

The November 7th ODS review does not adequately consider the substantial medical burden imposed on patients by rheumatoid arthritis: a chronic disease that can result in joint destruction, deformity, disability, and even premature death.[60] Rheumatoid arthritis is a painful, progressive disease without a cure. Moreover, the review does not adequately consider the considerable toxicities of other treatments for rheumatoid arthritis,[61] or that management of rheumatoid arthritis often requires that a range of treatment options be available. For example, treatment options are desirable because a patient may not respond adequately to, may not tolerate the toxicities of, and/or may not comply with other treatments.[62] Adequate consideration of such factors is necessary to make an appropriate, evidence-based, regulatory action that is in the best interest of the public health.

In short, an appropriate regulatory action should consider not only the risks of leflunomide but should also consider the risks of, and suffering caused by, rheumatoid arthritis (some of which can be alleviated by leflunomide). In addition, an appropriate regulatory action should take into account the comparative risks associated with other therapies for rheumatoid arthritis. After evaluation of such factors, as well as the risks associated with leflunomide, the full range of risk-management options for the drug should be considered.

[60] American College of Rheumatology Subcommittee on Rheumatoid Arthritis Guidelines. Guidelines for the management of rheumatoid arthritis; 2002 Update. Arthritis & Rheumatism 2002;46(2):328-346.
[61] Ibid. For example, see Table 5, pages 334-335.
[62] Ibid. For example, see Figure 1, page 329.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

Victor Raczkowski
2/4/03 04:08:32 PM
MEDICAL OFFICER
ODS Office Director memorandum on Arava (leflunomide)