Leflunomide Use and Risk of Lung Disease in Rheumatoid Arthritis: A Systematic Literature Review and Metaanalysis of Randomized Controlled Trials

Richard Conway, Candice Low, Robert J. Coughlan, Martin J. O’Donnell, and John J. Carey

ABSTRACT. Objective. To evaluate the relative risk (RR) of pulmonary disease among patients with rheumatoid arthritis (RA) treated with leflunomide (LEF).

Methods. We searched PubMed, Embase, and the Cochrane Central Register of Controlled Trials to April 15, 2014. We included double-blind randomized controlled trials (RCT) of LEF versus placebo or active comparator agents in adults with RA. Studies with fewer than 50 subjects or shorter than 12 weeks were excluded. Two investigators independently searched both databases. All authors reviewed selected studies. We compared RR differences using the Mantel-Haenszel random-effects method to assess total respiratory adverse events, infectious respiratory adverse events, noninfectious respiratory adverse events, interstitial lung disease, and death.

Results. Our literature search returned 5673 results. A total of 8 studies, 4 with placebo comparators, met our inclusion criteria. There were 708 respiratory adverse events documented in 4579 participants. Six cases of pneumonitis occurred, all in the comparator group. Four pulmonary deaths were reported, none in the LEF group. LEF was not associated with an increased risk of total adverse respiratory events (RR 0.99, 95% CI 0.56–1.78) or infectious respiratory adverse events (RR 1.02, 95% CI 0.58–1.82). LEF was associated with a decreased risk of noninfectious respiratory adverse events (RR 0.64, 95% CI 0.41–0.97).

Conclusion. Our study found no evidence of increased respiratory adverse events in RCT of LEF treatment. (First Release March 15 2016; J Rheumatol 2016;43:855–60; doi:10.3899/jrheum.150674)

Key Indexing Terms:
LEFLUNOMIDE METAANALYSIS RHEUMATOID ARTHRITIS
INTERSTITIAL LUNG DISEASE PULMONARY FIBROSIS

Rheumatoid arthritis (RA) is a chronic debilitating inflammatory joint disease affecting up to 1% of the population in developed countries. It is associated with significant symptoms, functional limitations, and an increased risk of mortality because of cardiovascular and other causes. Adequate treatment results in symptom improvement, normalization of physical and social functioning, and improvements in mortality. A variety of effective treatments exist for RA; however, patient and physician concerns regarding possible adverse events may limit optimal use in clinical practice.

A concern over the risk of pulmonary toxicity, in particular interstitial lung disease (ILD), has been raised for many of the drugs used to treat RA. Historically, methotrexate (MTX) was most commonly attributed as a causative agent in ILD. The introduction of leflunomide (LEF) to clinical practice was swiftly followed by reports of ILD associated with its use. Initial reports from Japan were subsequently supplemented by cases in Europe and the United States. Cases of ILD have also been attributed to the majority of biologic agents used to treat RA. ILD is an intrinsic part of the disease process in RA. Pulmonary complications in RA may also occur because of infections and unrelated pulmonary disease processes. Therefore, cases of ILD should be expected in patients with RA and it is important to compare any suspected drug-related risk with an appropriate control population.

Two systematic literature reviews of case reports and observational studies reported an increase in the risk of ILD in patients treated with LEF. Similar to any form of scientific work, observational studies have limitations, the most relevant of which are a susceptibility to ascertainment and channeling biases. Double-blind randomized controlled
trials (RCT) afford a setting that should limit these risks. A reported metaanalysis of double-blind RCT in patients with RA treated with MTX found an increased risk of infectious but not noninfectious pulmonary disease with MTX use, casting doubt on widely held beliefs on the drug’s pulmonary effects. The aim of our study was to evaluate the risk of ILD associated with LEF use by performing a metaanalysis of double-blind RCT in patients with RA.

MATERIALS AND METHODS
Data sources and searches. A systematic literature search was performed with no date limits using PubMed, Embase, and the Cochrane Central Register of Controlled Trials databases. The search was performed to April 15, 2014. We also searched for previously published metaanalyses and systematic literature reviews. The reference lists of relevant articles were reviewed. Full details of the search terms used are given in the supplementary appendix (available from the authors on request).

Study selection. The literature search was performed independently by 2 authors (CL and RC); discrepancies were resolved by consensus. The inclusion criteria for study selection were (1) double-blind RCT; (2) human studies; (3) patients with RA; (4) studies in English; (5) studies consisting of a minimum of 2 arms, at least 1 receiving LEF and at least 1 not receiving LEF; (6) studies including only adults (> 18 yrs); (7) trials of ≥ 12 weeks’ duration; (8) studies of ≥ 50 patients; and (9) studies reporting respiratory side effects for LEF and comparator groups separately. In the case of multiple publications of 1 RCT, we included the publication most relevant to our inclusion criteria involving detailed reporting of respiratory side effects. If the results of a study were reported at multiple timepoints, we included the publication of greatest duration, provided it remained a double-blind RCT and fully reported respiratory adverse events. If required, we reviewed previous publications of the same trial to fully assess the trial protocol and risk of bias. This approach was taken to avoid including study subjects more than once in the metaanalysis.

Relevant articles were selected using a 2-step approach. First, titles and abstracts of identified references were screened to exclude articles that did not deal with the topic of interest. Second, the full text of relevant articles was reviewed.

Data extraction and quality assessment. For each included study, data were extracted by 2 of the authors independently (CL and RC). Any discrepancies were resolved by discussion.

The data were entered into a database (RC) and checked by the remaining authors.

The following variables were extracted: authors, year of publication, population studied, number of patients, mean age and range, sex, disease duration, percentage of LEF-naive, previous immunosuppressive drug use, steroid use at baseline, study design and duration, comparator drug, and adverse events. Adverse events were extracted as both total respiratory and noninfectious adverse events. If the results of a study were reported at multiple timepoints, the study was reviewed. For each included study, data were extracted as both total respiratory and noninfectious adverse events relative to comparator agents (RR 0.99, 95% CI 0.56–1.78, I² = 88%; Figure 2). LEF was not associated with an increased risk of infectious adverse respiratory events (RR 1.02, 95% CI 0.58–1.82, I² = 88%; Supplementary Figure 1 is available from the authors on request). LEF was associated with a decreased risk of noninfectious respiratory adverse events (RR 0.64, 95% CI 0.41–0.97, I² = 0%; Supplementary Figure 2 is available from the authors on request). There were 6 reported cases of pneumonitis, all in patients treated with MTX in the comparator group (RR 0.20, 95% CI 0.02–1.75); no definitive diagnostic features or additional information were reported in these cases. There were 4 pulmonary deaths, all in patients treated with MTX in the comparator group (RR 0.26, 95% CI 0.03–2.20).

Risk of bias in included studies. In general, the data suggested a low risk of bias in the included studies (Supplementary Figure 3 and Supplementary Figure 4 are available from the authors on request). Inadequate information to assess the risk of bias in the included studies (Supplementary Figure 3 and Supplementary Figure 4 are available from the authors on request).
The funnel plot for total respiratory events showed no evidence of publication bias (Supplementary Figure 5 is available from the authors on request).

**Sensitivity analysis.** No significant effect of study size on overall results was found (< 400 patients, RR 0.79, 95% CI 0.58–1.09; ≥ 400 patients, RR 2.16, 95% CI 0.49–9.49). No difference in total respiratory adverse events was seen between studies with placebo (RR 0.93, 95% CI 0.60–1.44) compared with all active comparators (RR 1.10, 95% CI 0.50–2.45) or MTX (RR 0.91, 95% CI 0.82–1.02). While overall LEF was associated with a decreased risk of noninfectious respiratory adverse events, analysis by comparator agent did not reveal any differences (placebo, RR 1.03, 95% CI 0.04–25.00; all active comparators, RR 0.41, 95% CI 0.08–2.21; MTX, RR 0.41, 95% CI 0.08–2.21).

Analysis of study heterogeneity revealed that significant heterogeneity was introduced by the inclusion of the study by Chen, *et al*; we therefore repeated the analysis excluding the study20. This significantly reduced the $I^2$, but did not materially change the overall results. In our analysis, LEF was not associated with an increase in total respiratory adverse events (RR 0.90, 95% CI 0.81–1.01, $I^2 = 0%$) or infectious respiratory adverse events (RR 0.96, 95% CI 0.84–1.10, $I^2 = 0%$), and was associated with a decrease in noninfectious respiratory adverse events (RR 0.64, 95% CI 0.41–0.97, $I^2 = 0%$).

Exclusion of the studies reporting safety data on study completers only did not alter the overall results (RR for total respiratory adverse events 0.90, 95% CI 0.81–1.00, $I^2 = 0%$; RR for infectious respiratory adverse events 0.93, 95% CI 0.78–1.11, $I^2 = 14%$; RR for noninfectious respiratory adverse events 0.64, 95% CI 0.41–0.97, $I^2 = 0%$).

**DISCUSSION**

There was no increase in lung disease in patients with RA treated with LEF in our study. In prespecified subgroup analyses, LEF had a neutral effect on infectious pulmonary adverse events and decreased the risk of noninfectious pulmonary adverse events. The majority of studies used MTX as a comparator agent; previous work from our group demonstrated no association between MTX use and an increased risk of noninfectious pulmonary adverse events compared with other biologic and synthetic DMARD17. In our study, 6 reported cases of pneumonitis occurred in the MTX group compared with none in the LEF group; this did not reach statistical significance. Potential explanations include chance, a protective effect of LEF against RA-related ILD, a detrimental effect of MTX compared with LEF but not to other DMARD, or a small increased risk with MTX use, which falls within the CI of the previous studies17,28.

LEF was originally introduced in the United States and Europe at the end of the 1990s with no reports of a significant increase in pulmonary disease. However, soon after the introduction of LEF to Japan, case reports of ILD appeared10,11.
Subsequently, cases were also reported from the United States and Europe. Two systematic literature reviews of observational studies and case reports reported on LEF-related ILD. Roubille and Haraoui reported on 34 cases of LEF-induced or -exacerbated ILD, and while acknowledging the limitations of the data, concluded that LEF can be associated with rapid onset, potentially fatal ILD. Raj and Nugent reported on 138 cases of LEF-related ILD and concluded that LEF can cause ILD most commonly within the first 3 months of starting treatment. In contrast to these studies, a nested case-control study performed by Suissa, et al. concluded that the 2-fold increase in ILD with LEF use was likely the result of channeling bias. The disparity in results and conclusions present in the literature represents the ongoing uncertainty regarding LEF’s involvement in lung disease in clinical practice. We believe our study is concordant with the findings of Suissa, et al’s and can help ameliorate physician concern over LEF usage. All study designs have inherent strengths and weaknesses; the design of double-blind RCT intends to remove both overt and covert bias as confounding factors in study results. A significant limitation of the majority of randomized trials is that the study population is of relatively small size, reducing the power to detect small but clinically significant effects. Metaanalytical techniques help alleviate this limitation because the combination of studies increases the overall power, in addition to equally distributing inherent bias between the groups.

Similar to any study, ours also has some important limitations. Our study used previously performed clinical trials, therefore we were dependent on accurate recording and reporting of adverse events by the investigators involved in the original studies. As with previous work in this area, analysis was restricted to the adverse events detailed in the respective studies and we did not have access to the unpublished patient-level data. There was significant variation between studies in the frequency and terminology used to report adverse events. This is likely reflective of the adverse event definitions and reporting requirements of the original randomized trials. Included studies were of relatively short duration with a mean duration of 55 weeks; however, previous studies have reported that the majority of cases of purported LEF-related ILD occurred within the first 20 weeks of treatment. Participants in clinical trials are by design a more homogeneous and healthier population with fewer comorbidities; the results obtained from an evaluation of these trials may lack external validity. In particular, patients with preexisting lung disease are generally excluded from participation in clinical trials, limiting the generalizability of these findings to that particular population. A final limitation...

### Table 1. Characteristics of studies included in the metaanalysis.

<table>
<thead>
<tr>
<th>Study</th>
<th>Yr</th>
<th>LEF, n</th>
<th>Comparator, n</th>
<th>Study Duration, Weeks</th>
<th>Age, Yrs*</th>
<th>Female, %</th>
<th>Disease Duration, Yrs*</th>
<th>LEF-naive, %</th>
<th>Comparator Drug</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chen, et al</td>
<td>2013</td>
<td>886</td>
<td>898</td>
<td>52</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>PAE + CCPI</td>
</tr>
<tr>
<td>Jaime-Hernandez, et al</td>
<td>2012</td>
<td>43</td>
<td>42</td>
<td>52</td>
<td>43</td>
<td>87</td>
<td>NA</td>
<td>2</td>
<td>98 MTX</td>
</tr>
<tr>
<td>Ishaq, et al</td>
<td>2011</td>
<td>91</td>
<td>89</td>
<td>52</td>
<td>58</td>
<td>71</td>
<td>4</td>
<td>NA</td>
<td>MTX</td>
</tr>
<tr>
<td>Kremer, et al</td>
<td>2002</td>
<td>130</td>
<td>133</td>
<td>24</td>
<td>56</td>
<td>78</td>
<td>11</td>
<td>100</td>
<td>PBO</td>
</tr>
<tr>
<td>Cohen, et al</td>
<td>2001</td>
<td>190</td>
<td>318</td>
<td>104</td>
<td>54</td>
<td>73</td>
<td>7</td>
<td>NA</td>
<td>MTX PBO</td>
</tr>
<tr>
<td>Emery, et al</td>
<td>2000</td>
<td>501</td>
<td>498</td>
<td>104</td>
<td>58</td>
<td>71</td>
<td>4</td>
<td>NA</td>
<td>MTX</td>
</tr>
<tr>
<td>Smolen, et al</td>
<td>1995</td>
<td>133</td>
<td>225</td>
<td>24</td>
<td>59</td>
<td>73</td>
<td>7</td>
<td>100</td>
<td>SSZ PBO</td>
</tr>
<tr>
<td>Mladenovic, et al</td>
<td>1995</td>
<td>300</td>
<td>102</td>
<td>24</td>
<td>51</td>
<td>83</td>
<td>8</td>
<td>100</td>
<td>PBO</td>
</tr>
</tbody>
</table>

* Age and duration of disease expressed as mean as reported by individual studies. LEF: leflunomide; NA: not available; PAE: paeoniflorin; CCPI: cervus and cucumis polypeptide injection; MTX: methotrexate; SSZ: sulfasalazine; PBO: placebo.

![Figure 2](https://example.com/figure2.png)

**Figure 2.** Forest plot of relative risk for total adverse respiratory events for leflunomide compared with comparator agents. M-H: Mantel-Haenszel test; df: degrees of freedom.
may be considered the relatively low number of studies included in our metaanalysis. We elected to focus on large-scale high-quality studies, which included full reporting of the relevant data to minimize the effect of data collection errors on our results.

Our study has important clinical implications. The lack of evidence for an effect of LEF on lung disease suggests that previous reports may represent an association rather than a direct causative effect. Many factors including channeling bias and a genuine desire to report potential safety concerns may have contributed to this. While not completely excluding an effect of LEF on pulmonary disease, our data suggest that the extreme caution afforded this agent in patients with pulmonary disease may not be justified. Coupled with previous data on MTX, our data support increased use of pulmonary disease may not be justified. Coupled with previous data on MTX, our data support increased use of effective disease-modifying agents in the treatment of patients with RA, including those with coexistent or new-onset pulmonary disease. More severe RA is a risk factor for RA-related ILD, and the attainment of low disease activity with DMARD, including LEF, may contribute to a protective pulmonary effect.

The results of our metaanalysis demonstrate no increase in respiratory adverse events in patients with RA treated with LEF in double-blind RCT. Studies of pulmonary adverse events in patients treated with LEF and related agents for other diseases may provide further valuable information.

ACKNOWLEDGMENT
The PRISMA Checklist and other supplementary material are available from the authors on request.

REFERENCES

