The Comparative Efficacy and Safety of Biologics for the Treatment of Rheumatoid Arthritis: A Systematic Review and Metaanalysis

GERALD GARTLEHNER, RICHARD A. HANSEN, BETH L. JONAS, PATRICIA THIEDA, and KATHLEEN N. LOHR

ABSTRACT. Objective. Biologics are an important therapeutic option for treating patients with rheumatoid arthritis (RA). However, they are associated with rare but severe adverse events such as serious infections, lymphoma, or chronic heart failure. In addition, dosing regimens and routes of administration differ substantially among biologics. In a systematic review, we assessed the comparative efficacy and safety of biologic agents for RA.

Methods. We searched electronic databases up to May 2006. We limited evidence to controlled trials for efficacy but included observational evidence for safety. Outcomes of interest were clinical response, radiographic progression, and quality of life. Given the paucity of head-to-head evidence, we conducted adjusted, indirect comparisons of placebo-controlled trials.

Results. Twenty-six controlled trials provided efficacy data; 18 additional studies assessed safety. The only evidence directly comparing 2 biologic agents was a nonrandomized, open-label trial that found no differences in effectiveness and safety between etanercept and infliximab. Adjusted indirect comparisons indicate no significant differences in efficacy between anti-tumor necrosis factor (TNF) drugs. However, anti-TNF drugs appear to be more efficacious than anakinra, although not all comparisons reached statistical significance. Because of the lack of sound longterm safety data, evidence is insufficient to draw firm conclusions about the comparative safety of biologics.

Conclusion. Anti-TNF drugs appear to be more efficacious than anakinra but do not differ significantly among each other. Clinical considerations such as comorbidities, route of administration, dosing regimens, and specific side effect profiles may guide the choice of an anti-TNF drug. (First Release Nov 1 2006; J Rheumatol 2006;33:2398–408)

Key Indexing Terms:
RHEUMATOID ARTHRITIS          BIOLOGICS          TARGETED IMMUNE MODULATORS
SYSTEMATIC REVIEW               METAANALYSIS

Over the past decade, the treatment of rheumatoid arthritis (RA) has changed considerably with the advent of biologic agents such as abatacept, adalimumab, anakinra, etanercept, infliximab, and rituximab. Traditional disease modifying antirheumatic drugs (DMARD), in particular methotrexate (MTX), are still the cornerstone of most RA treatment regimens. However, toxicity may limit their use, and many patients do not respond adequately to traditional DMARD therapy. Thus, in patients with persistent disease despite aggressive management with oral agents, biologics, often in combination with MTX, are now considered the standard of care.

Biologics work by selectively blocking the effects of cytokines. For example, tumor necrosis factor (TNF) inhibitors, e.g., adalimumab, etanercept, infliximab, produce their primary effect by blocking the interaction of TNF-α with cell-surface receptors. Anakinra blocks interleukin 1 (IL-1), another naturally occurring cytokine. Abatacept inhibits T lymphocyte activation by binding to CD80 and CD86, thereby blocking interactions with CD28. Rituximab binds specifically to the antigen CD20, resulting in the depletion of B cells. All these actions greatly reduce various inflammatory and immunological responses.

Biologics differ considerably in dosing regimens and routes of administration. Abatacept, infliximab, and rituximab require intravenous administration. Abatacept infusions are
and infiximab infusions are repeated every 4 to 8 weeks, while rituximab is given at weekly intervals for a course of 4 to 8 weeks; adalimumab, anakinra, and etanercept can be administered subcutaneously by the patient. Administration intervals also differ substantially: adalimumab is administered once every other week, etanercept once a week, and anakinra daily. Table 1 summarizes biologics currently approved for the treatment of RA in the United States, including trade names, manufacturers, routes of administration, therapeutic mechanisms of action, and approved (labeled) uses.

To date, no head-to-head, double-blinded controlled trials (RCT) comparing one biologic to another have been published. Five metaanalyses of RCT provide good evidence about the general efficacy of anti-TNF drugs and anakinra for treating patients with RA. Only one systematic review conducted indirect comparisons of anti-TNF drugs, but this study included only 4 RCT. None of these meta-analyses included observational studies to document adverse events. In the case of biologics, RCT are compromised by small sample sizes and limited study durations to determine reliably rare but potentially fatal adverse events such as serious infections, lymphoma, autoimmunity, heart failure, or hepatotoxicity. Thus, they cannot reliably assess the risk-benefit profiles of biologics for this condition.

Our objective was to systematically review the comparative efficacy and safety of biologic agents for the treatment of RA in patients who have failed to respond to traditional DMARD therapy. To our knowledge, this is the first systematic review to combine the evidence of RCT and observational studies to determine the comparative risk-benefit profiles of biologic agents. This study is part of a larger systematic review of biologics conducted for the Drug Effectiveness Review Project.

**MATERIALS AND METHODS**

**Literature search.** To identify relevant articles we searched Medline®, Embase, The Cochrane Library, and the International Pharmaceutical Abstracts from 1980 to 2006 (up to May 2006); we used either Medical Subject Headings (MeSH) as search terms when available or key words when appropriate. We combined terms for RA (“arthritis,” MeSH, “arthritis, rheumatoid” (MeSH)) and adverse events (“adverse events,” “harms,” “drug reactions,” “toxicity”) with a list of 4 specific biologics and their trade names (“abatacept,” “adalimumab,” “anakinra,” “etanercept,” “ infiximab,” “rituximab,” or their respective trade names) from manually searched reference lists of pertinent review articles and letters to the editor. In addition, we explored the Center for Drug Evaluation and Research (CDER) database to identify unpublished research submitted to the US Food and Drug Administration (FDA). Pharmaceutical manufacturers were invited to submit dossiers, including citations, as outlined by the Drug Effectiveness Review Project.

**Study selection.** Two persons independently reviewed abstracts and full-text articles. Records were considered for exclusion if they did not meet pre-established eligibility criteria for study design or duration, patient population, interventions, outcomes, and comparisons to medications outside our scope of interest. The outcome of interest was clinical improvement as measured on a variety of scales. These included the Disease Activity Score (DAS-28), criteria from the American College of Rheumatology (ACR 20; ACR 50; ACR 70; Paulus criteria), radiographic progression, functional capacity, and quality of life. We included controlled trials lasting at least 12 weeks to determine comparative efficacy. For adverse events we included both experimental and observational studies. We limited observational studies to those with large sample sizes (> 100 patients) that lasted at least 3 months and reported an included outcome. We designed and used a structured data abstraction form to ensure consistency in appraisal for each study.

**Data abstraction and quality assessment.** Trained reviewers abstracted data from each study and assigned an initial quality rating. We assessed the internal validity (quality) of trials based on predefined criteria developed by the US Preventive Services Task Force (ratings: good, fair, poor) and the National Health Service Center for Reviews and Dissemination. We also

---

**Table 1. Biologic agents approved for the treatment of RA.**

<table>
<thead>
<tr>
<th>Generic Name</th>
<th>US Trade Name</th>
<th>Manufacturer</th>
<th>Route and Dosing</th>
<th>Half-life</th>
<th>Onset of Action</th>
<th>Mechanism of Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept</td>
<td>Orencia®</td>
<td>Bristol-Myers Squibb</td>
<td>Intravenous: 500 mg to 1000 mg dosed by weight; repeat at 2 and 4 weeks and then every 4 weeks thereafter</td>
<td>8-25 days</td>
<td>≥ 15 days</td>
<td>Cytotoxic T lymphocyte antigen immunoglobulin (CTLA-4Ig)</td>
</tr>
<tr>
<td>Adalimumab</td>
<td>Humira®</td>
<td>Abbott</td>
<td>Subcutaneous: 40 mg every other week as subcutaneous injection; may increase to 40 mg per week</td>
<td>10–18 days</td>
<td>1–14 days</td>
<td>TNF inhibitor</td>
</tr>
<tr>
<td>Anakinra</td>
<td>Kineret®</td>
<td>Amgen</td>
<td>Subcutaneous: 100 mg daily as subcutaneous injection; dose should be decreased to 100 mg every other day in renal insufficiency</td>
<td>7–8 hours</td>
<td>7–21 days</td>
<td>IL-1 receptor antagonist</td>
</tr>
<tr>
<td>Etanercept</td>
<td>Enbrel®</td>
<td>Amgen/Wyeth Immunex Centocor</td>
<td>Subcutaneous: 50 mg per week given as 1 or 2 subcutaneous injections</td>
<td>4.8 days</td>
<td>1–28 days</td>
<td>TNF inhibitor</td>
</tr>
<tr>
<td>Infliximab</td>
<td>Remicade®</td>
<td>Biogen Idec and Genentech</td>
<td>Intravenous: 3 mg/kg infusion at 0, 2, and 6 weeks followed by maintenance every 8 weeks; thereafter; may increase to maximum of 10 mg/kg every 4 weeks</td>
<td>9.8 days</td>
<td>7–14 days</td>
<td>TNF inhibitor</td>
</tr>
<tr>
<td>Rituximab</td>
<td>Rituaxan®</td>
<td>Biogen Idec and Genentech</td>
<td>Intravenous: 375 mg/m² infusion once weekly for 4 to 8 weeks</td>
<td>3–4 days</td>
<td>21–288 days</td>
<td>Anti-CD-20a</td>
</tr>
</tbody>
</table>

---


www.jrheum.orgDownloaded on February 15, 2021 from www.jrheum.org
assessed external validity (generalizability), but this did not influence quality ratings.

Data synthesis. Because only limited head-to-head evidence on biologics was available, we conducted indirect comparisons of placebo-controlled trials (employing the method proposed by Bucher, et al11 and metaregression). Evidence suggests that indirect comparisons agree with head-to-head trials if component studies are similar and if treatment effects are expected to be consistent across patients in different trials12.

Using random effects models, we calculated the pooled relative risks of achieving an ACR 20 or ACR 50 response for each biologic relative to placebo. ACR responses are defined as improvements on the 20%, 50%, or 70% level in counts of tender or swollen joints, pain score, patients’ and physicians’ global activity score, Health Assessment Questionnaire-Disability Index (HAQ-DI), and erythrocyte sedimentation rate13. For one study14, we assumed that Paulus criteria15 are comparable to ACR criteria. To reduce potential heterogeneity, we limited these analyses to populations that had remained symptomatic despite MTX treatment (i.e., we excluded MTX-naive populations). Further, we limited included data to FDA approved dosage ranges to achieve better equivalency across drugs. Data were insufficient to conduct quantitative analyses on any other outcomes than ACR 20 and ACR 50.

For each metaanalysis we assessed heterogeneity using the I² statistic. We explored heterogeneity with metaregression. We assessed publication bias using funnel plots and Kendall’s tests. Given the small number of component studies, results of these tests must be viewed cautiously. All statistical analyses were conducted using StatsDirect, version 2.3.8, and Stata 9.1.

Role of funding source. The funding source had no role in the study design, data collection, data analysis, data interpretation, or writing of this report. The corresponding author (GG) had full access to all data in the study and had final responsibility for the decision to submit for publication.

RESULTS
We identified 1419 citations from searches and reviews of reference lists. Figure 1 depicts the results of the literature search, showing disposition of articles and individual exclusion criteria. Overall, we included 26 controlled studies for efficacy and 18 additional studies of both experimental and observational designs for adverse events assessment. No RCT was a head-to-head trial. Efficacy studies were conducted in narrowly defined populations often limited to less than 1 year of followup. The mean age of study participants was 53.4 years; the majority was female (76.3%) and Caucasian (88.5%). All efficacy studies, except a nonrandomized trial16, were funded by the pharmaceutical industry.

Head-to-head evidence. The only evidence directly comparing 2 biologic agents was a nonrandomized, open-label trial from Europe that assessed the longterm (2 years) effectiveness and safety of etanercept, infliximab, and the DMARD leflunomide16. This study can be characterized as an effectiveness trial, with high generalizability of results. Etanercept had significantly greater response rates than infliximab at 3 months (p < 0.02; data not shown) and 6 months (p < 0.05; data not shown); no differences existed after 1 year. Otherwise, no evidence directly comparing the efficacy and safety of one biologic to another could be found.

Figure 1. Results of the literature search and disposition of the articles. Numbers here differ from numbers of included studies because a single study can lead to multiple publications.
Indirect comparisons. Given the paucity of direct head-to-head evidence, we conducted indirect comparisons of randomized placebo-controlled trials. As stated above, we limited analyses to MTX-resistant populations. We pooled data from 5 studies on adalimumab (n = 2354), 5 on etanercept (n = 1151), 4 on infliximab (n = 704), and 3 on anakinra (n = 1039). Table 2 summarizes characteristics of these studies; Table 3 presents studies not included. Data were insufficient to conduct indirect comparisons on abatacept and rituximab.

Overall, results of indirect comparisons indicate that efficacy does not differ substantially among anti-TNF drugs (adalimumab, etanercept, and infliximab). Table 4 presents relative risks for improvements on ACR 20 and ACR 50 measures. Given the wide confidence intervals we cannot exclude clinically significant differences with certainty.

Point estimates of comparative ACR 20 and ACR 50 responses consistently favor adalimumab, etanercept, and infliximab over anakinra. With 2 exceptions, however, differences did not reach statistical significance. Indirect comparisons of infliximab and of anti-TNF drugs as a class compared to anakinra yielded a statistically significantly greater efficacy on ACR 20 [relative risk 0.58 (95% CI 0.38–0.90) and RR 0.61 (95% CI 0.39–0.96), respectively], but not ACR 50.

Figures 2A and 2B depict ACR 20 and ACR 50 comparisons of anakinra with anti-TNF drugs as a class and individually. Sensitivity analyses (based on different study durations, concomitant MTX treatment, and disease durations) did not change the overall conclusions.

Few studies assessed long-term radiographic outcomes. In general, the rate of radiographic progression (e.g., Sharp score, joint erosions, joint space narrowing) was significantly lower in patients treated with biologics than in placebo-treated patients, regardless of concomitant DMARD therapy. Similarly, quality of life improved significantly for patients treated with biologics. Reported data for radiographic outcomes and quality of life, however, were insufficient for indirect comparisons.

Adverse events
Most studies that examined the efficacy of biologics also determined how well patients tolerated them. Some RCT had an open-label extension phase of up to 2 years, but their methods of adverse event assessment differed greatly. Few studies used objective scales such as the adverse reaction terminology from the World Health Organization (WHO). Most trials combined patient-reported adverse events with a regular

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>MTX</th>
<th>Mean Disease Duration, yrs</th>
<th>Radiographic Outcomes (biologic vs placebo)</th>
<th>ACR Response (%) of FDA Approved Doses Compared to Placebo</th>
<th>Quality Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furst 2003&lt;sup&gt;17&lt;/sup&gt;</td>
<td>636</td>
<td>24</td>
<td>ADA Yes 10.5</td>
<td>NR</td>
<td>53 vs 35 29 vs 11 15 vs 3</td>
<td>Fair</td>
</tr>
<tr>
<td>Keystone 2004&lt;sup&gt;18&lt;/sup&gt;</td>
<td>619</td>
<td>52</td>
<td>ADA Yes 11.0</td>
<td>Mean Sharp Score 0.45 vs 2.70</td>
<td>57 vs 24 40 vs 10 22 vs 5</td>
<td>Fair</td>
</tr>
<tr>
<td>Van de Putte 2003&lt;sup&gt;19&lt;/sup&gt;</td>
<td>284</td>
<td>12</td>
<td>ADA No 10.0</td>
<td>NR</td>
<td>50 vs 10 24 vs 1 11 vs 0</td>
<td>Fair</td>
</tr>
<tr>
<td>Van de Putte 2004&lt;sup&gt;20&lt;/sup&gt;</td>
<td>544</td>
<td>26</td>
<td>ADA No 11.0</td>
<td>NR</td>
<td>43 vs 19 21 vs 8 11 vs 2</td>
<td>Fair</td>
</tr>
<tr>
<td>Weinblatt 2003&lt;sup&gt;21&lt;/sup&gt;</td>
<td>271</td>
<td>24</td>
<td>ADA Yes 12.0</td>
<td>NR</td>
<td>67 vs 15 55 vs 8 27 vs 5</td>
<td>Fair</td>
</tr>
<tr>
<td>Klareskog 2004&lt;sup&gt;22&lt;/sup&gt;</td>
<td>682</td>
<td>52</td>
<td>ETA Yes 6.5</td>
<td>Mean Sharp Score –0.54 vs 2.80</td>
<td>85 vs 75 69 vs 43 40 vs 17</td>
<td>Good</td>
</tr>
<tr>
<td>Lan 2004&lt;sup&gt;23&lt;/sup&gt;</td>
<td>58</td>
<td>12</td>
<td>ETA Yes NR</td>
<td>NR</td>
<td>90 vs 34 66 vs 10 24 vs 0</td>
<td>Fair</td>
</tr>
<tr>
<td>Moreland 1999&lt;sup&gt;24-25&lt;/sup&gt;</td>
<td>234</td>
<td>12</td>
<td>ETA No 12</td>
<td>NR</td>
<td>59 vs 11 40 vs 5 15 vs 1</td>
<td>Fair</td>
</tr>
<tr>
<td>Moreland 1999&lt;sup&gt;26&lt;/sup&gt;</td>
<td>180*</td>
<td>12</td>
<td>ETA No NR</td>
<td>NR</td>
<td>75 vs 14 57 vs 7 NR</td>
<td>Fair</td>
</tr>
<tr>
<td>Weinblatt 1999&lt;sup&gt;27&lt;/sup&gt;</td>
<td>89</td>
<td>24</td>
<td>ETA Yes 13</td>
<td>NR</td>
<td>71 vs 27 39 vs 3 15 vs 0</td>
<td>Fair</td>
</tr>
<tr>
<td>Abe 2006&lt;sup&gt;28&lt;/sup&gt;</td>
<td>147</td>
<td>14</td>
<td>INF Yes 7.9</td>
<td>NR</td>
<td>57 vs 23 33 vs 4 13 vs 0</td>
<td>Fair</td>
</tr>
<tr>
<td>Kavanaugh 2000&lt;sup&gt;29&lt;/sup&gt;</td>
<td>28*</td>
<td>12</td>
<td>INF Yes 6.2</td>
<td>NR</td>
<td>50 vs 14 21 vs 14 NR</td>
<td>Fair</td>
</tr>
<tr>
<td>Maini 1999&lt;sup&gt;30,31&lt;/sup&gt;</td>
<td>428</td>
<td>30</td>
<td>INF Yes 8.4</td>
<td>NR</td>
<td>52 vs 17 33 vs 8 18 vs 2</td>
<td>Fair</td>
</tr>
<tr>
<td>Maini 1998&lt;sup&gt;14&lt;/sup&gt;</td>
<td>101*</td>
<td>26</td>
<td>INF Yes 10.0</td>
<td>NR</td>
<td>52 vs 7 47 vs 4 NR</td>
<td>Fair</td>
</tr>
<tr>
<td>Bresnihan 1998&lt;sup&gt;32&lt;/sup&gt;</td>
<td>472*</td>
<td>24</td>
<td>ANA No 3.9</td>
<td>NR</td>
<td>38 vs 27 15 vs 8 1 vs 1</td>
<td>Fair</td>
</tr>
<tr>
<td>Cohen 2002&lt;sup&gt;33&lt;/sup&gt;</td>
<td>419</td>
<td>24</td>
<td>ANA Yes 7.1</td>
<td>NR</td>
<td>42 vs 15 19 vs 3 7 vs 0</td>
<td>Fair</td>
</tr>
<tr>
<td>Cohen 2004&lt;sup&gt;34&lt;/sup&gt;</td>
<td>501</td>
<td>24</td>
<td>ANA Yes 10.5</td>
<td>NR</td>
<td>38 vs 22 17 vs 8 6 vs 2</td>
<td>Fair</td>
</tr>
</tbody>
</table>

* We included only results of FDA-approved dosing ranges. ADA: adalimumab, INF: infliximab, ANA: anakinra, MTX: methotrexate, ETA: etanercept. NR: not reported.
Clinical examination by an investigator. Determining whether assessment methods were unbiased and adequate was often difficult. Adverse events were rarely prespecified and defined. Only 2 RCT were designed to assess adverse events as primary outcomes17,50-52.

In addition to efficacy trials, we included 18 studies of both experimental and observational designs for adverse event assessment (Table 5). Many observational studies derived findings from the MedWatch adverse events reporting system of the FDA. It relies on voluntary reporting of adverse events, and underreporting is likely53. In addition, an adequate denominator to draw inferences about causation and the comparative risks of any drugs is lacking. Because data were insufficient to pool, we summarized the evidence qualitatively.

General tolerability. Overall, in efficacy trials, biologics appeared to have a good tolerability profile; rare but serious adverse events such as infections, lymphoma, or neutropenia were of concern but could not be assessed reliably in trials50-52,56,70,71. Discontinuation rates because of adverse events in patients treated with biologics ranged from 3% to 16%, and generally did not differ significantly from those in patients treated with placebo. Table 6 summarizes the adverse events most commonly reported in clinical trials.

The only head-to-head efficacy study16 also assessed differences in tolerability and safety between etanercept and infliximab, using the WHO adverse reaction terminology. Overall, etanercept and infliximab did not differ significantly in adverse events reported.

Injection site reactions (adalimumab, etanercept, anakinra) and infusion reactions (abatacept, infliximab, rituximab) were the most commonly and consistently reported adverse events. Injection site reactions were mainly erythema, pruritus, rash, and pain of mild to moderate severity. Nevertheless, these reactions were the most common reason for discontinuation attributable to adverse events. The mean, crude incidence of injection site reactions in RCT reviewed for this study was 19.0% (95% CI 9.2–28.8) for adalimumab, 25.0% (95% CI 11.2–38.1) for etanercept, and 55.8% (95% CI 4.9–100) for anakinra. The higher incidence of injection site reactions for anakinra than for adalimumab and etanercept is consistent with data reported in the respective package inserts72-74.

Some infusion reactions appeared to be more serious than injection site reactions. In clinical trials of infliximab, 17% of patients experienced infusion reactions consisting of mostly non-specific symptoms such as headache, dizziness, nausea, pruritus, chills, or fever56. However, 0.5% of patients had severe acute reactions that resembled acute anaphylactic conditions or led to convulsions56. In the open-label effectiveness study, 3.7% of patients treated with infliximab had a severe infusion reaction16; in a case series of 165 consecutive patients receiving infliximab this number was 1.0%54. Nevertheless, less than 2% of patients in clinical trials discontinued because of infusion reactions.

The rituximab studies reported infusion reactions in 35% to 45% of patients40-41: glucocorticoid premedication reduced this rate to 24%41. Infusion reactions in abatacept studies ranged from 5%43 to 29%46.

Specific adverse events. Serious infections. In efficacy trials, the incidence of serious infections was consistently higher in biologic- than in placebo-treated patients. However, clinically significant differences rarely reached statistical significance. For example, in a large safety RCT (n = 1414), a trend toward an increased risk of serious infections in anakinra-treated patients was apparent during the 6 months of treatment (2.1% anakinra vs 0.4% placebo; p = 0.068), but was not statistically significant50-52. A recent metaanalysis, pooling data of adalimumab and infliximab RCT, reported a 2-fold increase of serious infections (i.e., infections that required antimicrobial therapy and/or hospitalization) among patients treated with
anti-TNF drugs compared with those taking placebo (OR 2.0, 95% CI 1.3–3.1)\(^7\).

Long-term observational studies limited to assessments of infliximab and etanercept\(^56,70,76\) support these findings. The most common serious infection was tuberculosis\(^59\). A safety analysis of a Spanish registry of RA patients reported a more than 50 times greater risk of tuberculosis (RR 53.0, 95% CI 34.5–89.0) for RA patients treated with infliximab than for RA patients who did not receive anti-TNF drugs\(^58\).

Several observational studies indicate that infliximab might lead to a higher risk of tuberculosis or other granulomatous infections, and may lead to a faster outbreak of tuberculosis than etanercept\(^57,59-63\). In one study, the median interval from start of infliximab therapy to diagnosis of tuberculosis was 3 months\(^59\). By contrast, a different analysis of safety data, published in abstract form only, concerning etanercept and tuberculosis reported a median time of 11.5 months from start of etanercept therapy to diagnosis of tuberculosis\(^77\).
Some of these findings, however, were derived from the MedWatch spontaneous reporting system of the FDA and must be interpreted cautiously. Nevertheless, the higher incidence of granulomatous infections in infliximab-treated patients is consistent across multiple studies.

**Lymphoma.** The risk of both Hodgkin’s disease and non-Hodgkin’s lymphoma is generally increased in patients with RA. Data from controlled trials do not provide sufficient evidence concerning the comparative risk attributable to either biologics or a combination of biologics and MTX.

A large prospective cohort study followed 18,572 patients with RA registered in the National Data Bank for Rheumatic Diseases.
Diseases for up to 3 years. Lymphomas were more common in patients undergoing anti-TNF therapies, but confidence intervals for treatment groups overlapped. MedWatch identified 26 reported cases of lymphoproliferative disorders in patients treated with infliximab or etanercept for Crohn’s disease or RA as of 2006. In some cases, lymphoma developed shortly after starting therapy; regression occurred in 2 patients after discontinuing therapy.

A recent meta-analysis pooled data on malignancy rates in efficacy trials of adalimumab and infliximab. The results presented a more than 3-fold increase of malignancies in patients treated with adalimumab or infliximab compared to those receiving placebo (OR 3.3, 95% CI 1.2–9.1). Existing evidence is insufficient to draw conclusions about an increased risk of specific malignancies other than lymphoma for patients receiving biologics. A clinical trial database review did not detect an increased incidence of squamous cell carcinoma in 1442 RA patients (4257 patient-years) treated with etanercept (crude rate 2.8 cases per 1000 patients). However, the median follow-up time was only 3.7 years.

Congestive heart failure. A MedWatch analysis reported that half the patients who developed new-onset congestive heart failure (CHF) under etanercept or infliximab treatment had no identifiable risk factors. No direct evidence on the comparative risk of CHF exists. Indirect evidence comes from 3 trials, 2 on etanercept and one on infliximab, that evaluated the efficacy of these drugs for the treatment of CHF. Study populations had no rheumatic diseases. At least one etanercept study presented an increased risk of worsening heart failure. Similarly, the infliximab study presented higher mortality rates in the 10 mg/kg arm than in the placebo and 5 mg/kg arms. The infliximab package insert contains contraindication for use in patients with CHF; the package inserts for etanercept and adalimumab emphasize caution.

Other adverse events. Evidence from randomized trials and observational studies is insufficient to draw conclusions regarding the risk of rare but serious adverse events such as demyelination, autoimmunity, neutropenia, and hepatotoxicity. A case series based on MedWatch data indicated that infliximab and etanercept might be associated with demyelination. Similar cases have been seen in regulatory trials of adalimumab. All neurologic events were partially or completely resolved after discontinuation of treatment.

Controlled trials and observational studies have not confirmed reports of autoimmunity. However, case reports suggest an association between infliximab and drug-induced lupus erythematosus and other autoimmune diseases. The infliximab package insert reports that 34% of patients treated with infliximab and MTX experienced transient elevations of liver function measures. Severe liver injury, including acute liver failure, has been reported. Hepatotoxicity has not been reported for other biologics.

DISCUSSION

In this systematic review, we combined clinical trial data of 17 placebo-controlled studies including 5248 patients with MTX-resistant RA. Indirect comparisons indicate that no substantial differences in efficacy exist among anti-TNF drugs. These findings are consistent with a nonrandomized, open-label effectiveness trial comparing etanercept to infliximab, the only direct evidence to date. Further, an earlier meta-analysis based on 4 trials reached a similar conclusion.

By contrast, results also indicate that anakinra is less efficacious than anti-TNF drugs. Although most comparisons do not reach statistical significance because of wide confidence intervals, a trend favoring anti-TNF drugs over anakinra is obvious. In addition, this finding is largely consistent with a meta-analysis and adjusted indirect comparisons conducted by the UK Health Technology Assessment Programme, which found anakinra to be less efficacious than anti-TNF drugs as a class (limited to infliximab and etanercept).

Data were insufficient to conduct indirect comparisons on abatacept and rituximab.

An important challenge for our systematic review was the lack of long-term studies with the methodological strength to assess rare but severe adverse events. Currently, no conclusions can be drawn regarding the comparative safety of biologics, although observational evidence indicates that some differences might exist. Biologics do not appear to differ substantially in short-term tolerability and safety. Differences in tolerability exist primarily with respect to adverse events caused by the route of administration. Anakinra has a substantially higher rate of injection site reactions than anti-TNF drugs. Abatacept, infliximab, and rituximab carry the risk of severe infusion reactions that cannot occur in drugs administered subcutaneously. Observational studies indicate possible differences among biologics with respect to serious infections, hepatotoxicity, or chronic heart failure. However, the evidence is weak and must be interpreted cautiously.

Our study has several limitations. First, indirect comparisons have methodological drawbacks and do not possess the validity of direct head-to-head trials. Further, because indirect comparisons are low in power, confidence intervals for all comparisons are wide and encompass differences that would be clinically significant. Although results of our adjusted indirect comparisons can be viewed as the best available comparative evidence, inferences must be drawn cautiously. Because of the methodological limitations, we did not attempt to calculate numbers needed to treat to illustrate differences in effect sizes.

Second, owing to limitations in reported data, we had to constrain our analyses to ACR 20 and ACR 50 response rates. Physicians need to keep in mind that, despite a clinical response, joint damage might progress. Radiographic and other measures of joint destruction and functional capacity would be necessary to assess the comparative efficacy on disease progression. Because biologics are relatively new agents,
longterm, controlled studies are generally lacking — a fact that also severely limits the comparative safety assessment. Existing data on rare but severe adverse events stem mainly from voluntary adverse event reporting systems, which cannot assess causation or the comparative safety of drugs. This lack of sufficient data severely compromises any assessment of the risk-benefit profile of biologic agents.

Third, most included efficacy trials were conducted in highly selected populations. Further, we limited our analyses to patients who failed traditional DMARD treatment. Therefore, our results may have limited generalizability and cannot be extrapolated to DMARD-naive patients with early disease.

Decisions about the choice of a biologic for the treatment of RA include not only efficacy and safety but also dosing regimens, routes of administration, comorbidities, costs, and insurance coverage. Given similar efficacy among anti-TNF drugs, such factors can guide clinical decisions.

An important contribution of any systematic review is its ability to highlight gaps in the scientific evidence. In the case of biologics, the lack of head-to-head and controlled longterm studies significantly limits our knowledge about their comparative risk-benefit profiles. When high quality evidence is missing, “weaker” evidence such as case reports should be taken into consideration. Future research must clarify the remaining questions and clearly address the longterm safety issues of biologics.

ACKNOWLEDGMENT
We thank Timothy Carey, MD, MPH, and Marian McDonagh, PharmD, for insightful comments on earlier versions of this report. Many thanks also to Laura Morgan for help with database management and word processing.

REFERENCES
23. Lan JL, Chou SJ, Chen DY, Chen YH, Hsieh TY, Young MJ. A


75. Coletta AP, Clark AL, Banarjee P, Cleland JG. Clinical trials update: RENEWAL (RENAISSANCE and RECOVER) and ATTACH. Eur J Heart Fail 2002;4:559-61.


80. Coletta AP, Clark AL, Banarjee P, Cleland JG. Clinical trials update: RENEWAL (RENAISSANCE and RECOVER) and ATTACH. Eur J Heart Fail 2002;4:559-61.

