Comparative Effectiveness Research (CER): A Summary of AHRQ’s CER on Therapies for Rheumatoid Arthritis

Gary M. Oderda, PharmD, MPH, and Lisa M. Balfe, MPH

ABSTRACT

BACKGROUND: In recent years, the U.S. government has designated funding of several large-scale initiatives for comparative effectiveness research (CER) in health care. The American Recovery and Reinvestment Act (ARRA) of 2009 apportioned more than $1 billion to support CER programs administered by the Department of Health and Human Services (DHHS), the National Institutes of Health (NIH), and the Agency for Healthcare Research and Quality (AHRQ). CER is generally defined as the undertaking of original research or systematic reviews of published literature in order to compare the benefits and risks of different approaches to preventing, diagnosing, or treating diseases. These approaches may include diagnostic tests, medications, medical devices, and surgeries. The overall goals of CER are to support informed health care decisions by patients, clinicians, payers, and policy makers and to apply its evidence to ultimately improve the quality, effectiveness, and efficiency of health care.

OBJECTIVES: To (a) provide managed care professionals with general definitions of CER, specifically as it is administered by AHRQ; (b) discuss the importance of CER to clinical and managed care pharmacists; and (c) summarize key methods and findings from AHRQ’s 2007 comparative effectiveness review on therapies for rheumatoid arthritis (RA).

SUMMARY: As supported by AHRQ, CER is conducted in order to synthesize comprehensive evidence on the comparative benefits and harms of treatment interventions. The findings from comparative effectiveness reviews can thus contribute to informing therapeutic strategies and treatment decisions. In 2007, a multitude of RA treatment options and studies motivated AHRQ to commission a systematic comparative effectiveness review. Conducted by investigators at the RTI-University of North Carolina Evidence-Based Practice Center, the review included comparisons of synthetic disease-modifying antirheumatic drugs (DMARDs), biologic agents, synthetic DMARDs versus biologic agents, and various combination therapies. Head-to-head comparisons of synthetic DMARDs generally revealed no significant differences in long-term clinical and radiographic outcomes, or in functional capacity or health-related quality of life. Two nonrandomized prospective cohort studies and 1 open-label effectiveness trial reported no differences in ACR20 and ACR50 response rates in patients treated with the tissue necrosis factor (TNF)-alpha inhibitors etanercept and infliximab. Comparisons of TNF-alpha inhibitors generally indicated no significant differences in rates of adverse events, including serious infections, and no increases in rates over time. In comparisons of a biologic agent combined with methotrexate versus a biologic agent alone, combination therapies were generally associated with better clinical response rates and better outcomes of functional capacity and quality of life. The most common adverse events observed in studies on biologic agents were diarrhea, headache, nausea, rhinitis, injection site reactions, and upper respiratory tract infections.

J Manag Care Pharm. 2011;17(9-b):S19-S24

Copyright © 2011, Academy of Managed Care Pharmacy. All rights reserved.
comparisons will be made between interventions, instead of comparisons of interventions with placebo, which is common in randomized controlled trials (RCTs). The second component is that the universe being compared is much broader than just looking at drug therapy. The Institute of Medicine developed a list of the 100 highest priorities for CER. Of these, only 36 involved either direct comparisons of drug therapy or a comparison of drug therapy with other treatment options.

In looking at CER, one must distinguish effectiveness from efficacy. In 1962, Congress passed the Kefauver-Harris amendments requiring sponsors to demonstrate the safety and efficacy of a drug to be licensed for sale in the United States. Efficacy is proven through RCTs. These trials enroll relatively small numbers of subjects, hundreds or a few thousand, who meet rigorous inclusion and exclusion criteria related to the disease state of interest, comorbid conditions, and additional drug therapy. These trials demonstrate efficacy, or whether the drug can work in a very select group of patients. These studies lack the external validity to determine whether the drug actually works in a “real-world” group of patients who have multiple complicating disease states, take a variety of other drugs, and may or may not be adherent to therapy. Studies that look at “real-world” data measure effectiveness or how well the drug works in a broad group of patients outside of a clinical trial.

Does CER really mean just effectiveness, or will clinical trial data be used as well? At present this issue is not clear; however, an attractive alternative, called pragmatic clinical trials, has been described by Sullivan and Goldmann (2011), as well as others. Pragmatic clinical trials are prospective in nature but address the issue of external validity by enrolling the full range of patients that clinicians in practice will see as potential candidates to use the drug of interest. A component of CER will include primary studies which develop new evidence through original studies that make head-to-head comparisons of treatment alternatives for a condition or a disease. Additionally, an important option is to look at retrospective database studies for CER. These studies either use claims data, which include medical and pharmacy claims for a large population of patients, electronic medical record data, or both. Claims data have the advantage of being able to determine which drugs and treatments were actually used and when. However, claims data lack many meaningful clinical outcomes. Electronic medical record (EMR) data are an excellent source of clinical outcome information, including data on which drugs were prescribed. Nonetheless, EMR data do not generally indicate whether prescriptions were actually filled. The Holy Grail for these types of studies is a combination of both types of data, which characterizes some limited data sets in closed systems such as the U.S. Department of Veterans Affairs and Kaiser Permanente. There is a significant need for comparative effectiveness studies to investigate beyond the scope of placebo-controlled clinical trials used for licensing. For example, CER conducted as systematic reviews and meta-analyses of published studies affords the potential to include head-to-head trials, clinical trials that enroll a broader group of patients, and observational studies containing claims and EMR data.

AHFQ Methods for Conducting CER Systematic Reviews

Comparative effectiveness reviews are comprehensive reports that use published data to make head-to-head comparisons of products. These reviews are conducted by AHRQ’s 14 Evidence-Based Practice Centers (EPC). Established in 2007, the EPCs are listed as follows:

- Blue Cross and Blue Shield Association, Technology Evaluation Center
- Duke University
- ECRI Institute
- Johns Hopkins University
- McMaster University
- Minnesota Evidence-Based Practice Center
- Oregon Evidence-Based Practice Center
- RTI International—University of North Carolina
- Southern California
- Tufts—New England Medical Center
- University of Alberta
- University of Connecticut
- University of Ottawa
- Vanderbilt University

AHRQ has established 3 key principles for conducting comparative effectiveness reviews:

- Reviews must be relevant and timely for decision makers
- Reviews must be objective and scientifically rigorous
- Public participation in the reviews and transparency in the review process are critical to increase public confidence and credibility

AHRQ further points out that interpretation of the data is critical and that there are limits of interpretation. Drugs may be shown to be equivalent for an average group of patients, but that does not necessarily imply that they are equivalent for all individuals. As the EPCs conduct reviews, they are further guided by principles for the Effective Health Care Program, which include:

- Approaching the evidence from a clinical, patient-centered perspective
- Fully exploring the clinical logic underlying the rationale for a service
- Casting a broad net with respect to types of evidence, which includes placing a high value on effectiveness and applicability, in addition to internal validity
- Presenting benefits and harms for different treatments and tests in a consistent manner

As part of the transparency, topics are nominated in a public
Comparative Effectiveness Research (CER): A Summary of AHRQ’s CER on Therapies for Rheumatoid Arthritis

process. AHRQ contracts with 1 of the EPCs to complete each review. This process includes an expert panel of researchers and clinicians who guide the selection of key clinical questions for the topic. The EPC conducts an extensive literature search and selects studies according to rigorous inclusion criteria. For example, studies are generally screened for whether patient samples represent prespecified populations and conditions of interest, appropriateness of blinding methods, and sufficiency of patient follow-up. From the studies that meet the inclusion criteria, the EPC extracts data relevant to the key clinical questions, evaluates the strength of evidence, compiles the data and performs meta-analyses, and prepares a draft technical report that undergoes peer review, public commentary, and a revision process before the report is published on the Effective Health Care Program website.

Why Is CER Important for Pharmacists?
Therapeutic decisions are often made through anecdote and personal experience with individual patients. In many cases pharmacists need to make decisions based on limited information from placebo-controlled clinical trials when they really need head-to-head data. CER will provide pharmacists with evidenced-based data that compares therapeutic agents used on real-world patients. Comparisons of head-to-head trials will allow pharmacists to determine the most effective and safe drug for patients and to apply that information to an individual patient. The data can be used by pharmacists to make clinically informed decisions for patients and to communicate that information to other practitioners.

AHRQ’s 2007 CER on Rheumatoid Arthritis Therapies
In 2007, AHRQ published a systematic review of research on the comparative benefits and risks of synthetic and biologic disease-modifying antirheumatic drugs (DMARDs) for RA and psoriatic arthritis in adults. Conducted by the RTI-University of North Carolina EPC, the review included 42 head-to-head trials and 58 observational studies on RA therapies. Published studies were identified and obtained through comprehensive database searches covering the period from 1980 through September 2007. The key clinical questions that guided the review are summarized as follows:
1. Do RA therapies differ in their ability to reduce patient-reported symptoms, to slow or limit progression of radiographic joint damage, or to maintain remission?
2. Do RA therapies differ in their ability to improve functional capacity or quality of life?
3. Do RA therapies differ in harm, tolerability, adherence, or adverse effects?
4. What are the comparative benefits and harms of RA therapies in patient subgroups specified by stage of disease, history of prior therapy, demographics, concomitant therapies, or comorbidities?
In the studies that met the EPC researchers’ inclusion criteria, analyses involved various comparisons among and between synthetic DMARDs (corticosteroids, hydroxychloroquine, leflunomide, methotrexate, and sulfasalazine) and biologic agents (abatacept, adalimumab, anakinra, etanercept, infliximab, and rituximab) used as monotherapy or in combination therapy. The researchers evaluated the strength of study evidence using a modified version of the GRADE method (Grading of Recommendations, Assessment, Development, and Evaluation). The strength of evidence for each treatment comparison was graded as high, moderate, or low, reflecting the reviewers’ level of confidence in whether the evidence reflected the true effects of the study interventions.

Most of the studies included in the AHRQ systematic review evaluated clinical improvement using American College of Rheumatology (ACR) 20/50 criteria and disease activity scores (DAS). Radiographic progression was most often evaluated with Sharp or Sharp—van der Heijde scores. Functional capacity and health-related quality of life outcomes were evaluated with the Health Assessment Questionnaire (HAQ) or the Medical Outcomes Study Short Form 36 (SF-36). Analyses of comparative risks of RA therapies were based on 1 or more of the following adverse effects: hepatic events, interstitial lung disease, respiratory infection, hospitalization due to infection, and malignancies. In addition, the systematic review analyzed the comparative risks of serious infections, including tuberculosis, pneumonia, osteomyelitis, progressive multifocal leukoencephalopathy, and sepsis. Addressing key questions 1-3, notable findings regarding the comparative benefits and risks of RA therapies are summarized as follows.

Comparisons of Synthetic DMARDs Used as Monotherapy.
With few exceptions, head-to-head comparisons revealed no significant differences in long-term clinical and radiographic outcomes or in functional capacity or health-related quality of life among patients treated with various synthetic DMARDs used as monotherapy. Unless noted otherwise in the following summaries, rates of adverse events and discontinuations did not differ in comparisons of synthetic DMARDs.

Comparison of leflunomide versus methotrexate was based on 1 publication that reported a meta-analysis of 2 RCTs conducted over 2 years. After 1 year, the proportion of patients meeting ACR20 criteria was significantly greater in the methotrexate versus leflunomide arms (odds ratio [OR] = 1.43, 95% confidence interval [CI] = 1.15-1.77, P = 0.001). At the 2-year endpoint, there was no significant difference in ACR20 outcomes between the treatment groups (OR = 1.28, 95% CI = 0.98-1.67). However, for the 2-year measures of functional capacity and health-related quality of life, significantly less improvement was reported for patients using methotrexate versus leflunomide (SF-36 physical component: 4.6 vs. 7.6, P < 0.01; HAQ Disability Index: −0.26 vs. −0.45, P < 0.01). Radiographic outcomes did not differ significantly between the treatment
For all of the above comparisons, the EPC researchers judged the strength of evidence as moderate.

Based on 3 RCTs that lasted up to 1 year (N = 479), the EPC researchers concluded that there were no significant differences in clinical outcomes and functional capacity in patients treated with methotrexate versus sulfasalazine. For this comparison, a difference in treatment persistence was reported in 1 meta-analysis of 71 RCTs and 88 observational studies. After 5 years, a greater proportion of patients reported continuing methotrexate (36%) versus sulfasalazine (22%, P value not reported).

**Comparisons of Biologic Agents Used as Monotherapy.** In their literature search, the EPC researchers did not identify any RCTs with head-to-head comparisons of biologic agents. Two nonrandomized prospective cohort studies and 1 open-label effectiveness trial compared the tissue necrosis factor (TNF)-alpha inhibitors etanercept and infliximab. Based on ACR20 and ACR50 response rates, there were no differences in efficacy outcomes between treatment groups; the strength of evidence in these 3 studies was judged as moderate. As a class, tissue necrosis factor (TNF)-alpha inhibitors were significantly more effective than anakinra based on ACR20 response rates. Reported as incidence rates, the most common adverse events observed in observational studies on biologic agents were diarrhea (7%-18%), headache (12%-18%), nausea (8%-20%), rhinitis (8%-18%), injection site reactions (19%-56%), and upper respiratory tract infections (9%-24%). Comparisons of TNF-alpha inhibitors generally indicated no significant differences in rates of adverse events, including serious infections, and no increases in rates over time. A systematic review indicated higher rates of injection site reactions in patients using anakinra compared with adalimumab or etanercept. A retrospective cohort study reported higher overall discontinuation rates for anakinra (41%) compared with etanercept (31%; P = 0.004) or infliximab (35%; P = 0.03). Pooled results from 3 observational studies revealed that the incidence of granulomatous infections was higher in patients treated with infliximab (239 infections per 100,000 patients) versus etanercept (74 infections per 100,000 patients).

**Synthetic DMARDs Versus Biologic Agents Used as Monotherapy.** The EPC researchers identified 3 RCTs that compared the effectiveness of methotrexate versus the TNF-alpha inhibitors adalimumab or etanercept. In 2 of these trials, all patients had early RA and were methotrexate-naive. The other trial included methotrexate-naive patients as well as patients who had not responded to a synthetic DMARD other than methotrexate. In all 3 studies, there were no significant treatment-group differences in clinical response, quality of life, or functional capacity. However, radiographic outcomes were significantly better among patients treated with a biologic agent versus methotrexate. For example, as reported in the ERA (Early Rheumatoid Arthritis) study, over a 1-year follow-up period a lack of radiographic progression was observed in 72% of patients on etanercept versus 60% of patients on methotrexate (P = 0.007). In 1-year prospective cohort study, subjects who had failed initial RA treatment had a markedly higher probability of achieving functional independence on a biologic agent than on a synthetic DMARDs (OR = 3.88, 95% CI = 1.71-8.79). In addition, this study reported that patients using biologic agents were more likely to achieve remission (OR = 1.95; 95% CI = 1.20-3.19). Otherwise, the AHRQ systematic review indicated no differences in incidences of adverse events between biologic agents and synthetic DMARDs; however, in the studies on which comparisons were based, small sample sizes may have limited the statistical power necessary to detect differences.

**Monotherapy Versus Combination Therapy.** Two 4-year trials were identified that compared sulfasalazine-methotrexate to monotherapy with 1 of these medications in patients with early RA. The findings for both studies indicated no differences in ACR response rates or radiographic changes across the combination therapy versus monotherapy groups. However, over an 18-month period, another study reported a significantly greater improvement in DAS measures for patients treated with combination therapy (–0.67) versus methotrexate alone (–0.26, P = 0.023). In 3 RCTs, synthetic DMARD plus corticosteroid therapy was compared to synthetic DMARD monotherapy. These studies reported significantly fewer eroded joints and lower radiographic progression in patients on the combination therapy. In 1 study, patients on combination therapy (corticosteroid combined with sulfasalazine or methotrexate) had significantly higher remission rates than patients on DMARD monotherapy (DAS28 < 2.6: 55.5% vs. 43.8%, P < 0.001). Studies included in this analysis reported similar discontinuation rates due to adverse events. In the few trials that have compared combination therapy comprising 2 biologic agents with biologic monotherapy, the most common finding is that combination therapy is associated with higher rates of adverse events. In 1 study comparing etanercept alone versus a combination of etanercept and anakinra, the combination therapy group experienced significantly more adverse events (14.8% vs. 2.5%). Likewise, another study reported more adverse events among patients treated with combinations of abatacept plus various other biologic agents versus abatacept alone (22.3% vs. 11.7% to 12.5%, P value not reported).
One of the more consistent and robust findings of the AHRQ systematic review involved comparisons of a biologic agent combined with methotrexate versus a biologic agent alone. The combination therapies included methotrexate plus adalimumab, etanercept, infliximab, or rituximab. Compared with a biologic agent alone, combination therapy was generally associated with better clinical response rates. For example, 1 trial reported that significantly more patients on adalimumab plus methotrexate compared with adalimumab monotherapy met ACR50 criteria after 2 years of treatment (59% vs. 37%, \( P<0.001 \)). Similarly, the TEMPO (Trial of Etanercept and Methotrexate with Radiographic Patient Outcomes) study also identified higher DAS28 remission rates for the combination therapy compared with etanercept alone (35% vs. 16%, \( P<0.001 \)). Combination therapy was also associated with significantly better outcomes of functional capacity and quality of life.

### Conclusions

For managed care professionals, comparative effectiveness research offers value through its focus on making meaningful comparisons between health care interventions instead of comparing interventions with placebo, which is common in RCTs. In addition, comparative effectiveness reviews, like the AHRQ review on RA therapies, offer the advantage of pooling findings from numerous studies and grading the quality of evidence to guide treatment and management decisions. In December 2010, AHRQ published a protocol for conducting an update to its 2007 comparative effectiveness review on RA therapies. The updated review, to be based on the same key clinical questions that guided the original study, will evaluate outcomes of RCTs and observational studies that have been published since 2007. In addition to the synthetic DMARDs and biologic agents that were compared in the 2007 review, the update encompasses 3 biologics—certolizumab pegol, golimumab, and tocilizumab—that were approved in 2009 or 2010.

### DISCLOSURES

This supplement was sponsored by PRIME Education, Inc. and Purdue University through an independent educational grant from Genentech, Biogen Idec, and Centocor Ortho Biotech Services LLC. Gary Oderda received compensation from PRIME Education, Inc. for participating in the live continuing education activity on which this article is based and for writing parts of the article. Gary Oderda reports consulting and other financial relationships including advisory board membership with several companies with commercial products that may be used in rheumatoid arthritis including Abbott, Janssen, King Pharmaceuticals, Novartis, Pacira, and Pfizer.

Lisa Ballé wrote summaries of the findings from the comparative effectiveness studies reviewed in this article. PRIME Education, Inc. is a medical education company that receives grants and funding for educational programs from various pharmaceutical manufacturers.

### REFERENCES


### Authors

GARY M. ODERDA, PharmD, MPH, is Professor and Director of the University of Utah Pharmacotherapy Outcomes Research Center in Salt Lake City, Utah. LISA M. BALFE, MPH, is a medical writer at PRIME Education, Inc. in Tamarac, Florida.

AUTHOR CORRESPONDENCE: Gary M. Oderda, PharmD, MPH, University of Utah College of Pharmacy, 421 Wakara Way, Suite 208, Salt Lake City, UT 84108. Tel: 801.581.6257; E-mail: goderda@pharm.utah.edu.


