Drug Class Review
on
Targeted Immune Modulators

Final Report Update 1
January 2007

Original Report Date: December 2005
A literature scan of this topic is done periodically

The purpose of this report is to make available information regarding the comparative effectiveness and safety profiles of different drugs within pharmaceutical classes. Reports are not usage guidelines, nor should they be read as an endorsement of, or recommendation for, any particular drug, use or approach. Oregon Health & Science University does not recommend or endorse any guideline or recommendation developed by users of these reports.

Gerald Gartlehner, MD, MPH
Richard A. Hansen, PhD
Patricia Thieda, MA
Beth Jonas, MD
Kathleen N. Lohr, PhD
Tim Carey, MD, MPH

Produced by
RTI-UNC Evidence-based Practice Center
Cecil G. Sheps Center for Health Services Research
University of North Carolina at Chapel Hill
725 Airport Road, CB# 7590
Chapel Hill, NC 27599-7590
Tim Carey, MD, MPH, Director

Oregon Evidence-based Practice Center
Mark Helfand, MD, MPH, Director

Copyright © 2006 by Oregon Health & Science University
Portland, Oregon 97201. All rights reserved.
Note: A scan of the medical literature relating to the topic is done periodically (see http://www.ohsu.edu/ohsuedu/research/policycenter/DERP/about/methods.cfm for scanning process description). The Drug Effectiveness Review Project governance group elected to proceed with another update of this report. Please see timeline on the DERP website for details on the date of its release. Prior versions of this can be accessed at the DERP website.
IV. Psoriatic arthritis (PsA) ............................................................................................................................... 52
  A. Summary of the Evidence .......................................................................................................................... 52
  B. Description of Studies ............................................................................................................................... 52
  C. Study Populations ........................................................................................................................................ 52
  D. Outcome Measures ..................................................................................................................................... 52
  E. Methodological Quality ............................................................................................................................. 53
  F. Sponsorship ............................................................................................................................................... 53
  G. Comparative Efficacy and Effectiveness ................................................................................................. 53
  H. General Efficacy ....................................................................................................................................... 53
    Adalimumab .......................................................... 53
    Alefacept .............................................................. 54
    Etanercept ............................................................ 54
    Infliximab ............................................................. 55

V. Crohn's Disease ......................................................................................................................................... 57
  A. Summary of the evidence .......................................................................................................................... 57
  B. Description of Studies ............................................................................................................................... 57
  C. Study Populations ....................................................................................................................................... 58
  D. Outcome Measures ..................................................................................................................................... 58
  E. Methodological Quality ............................................................................................................................. 58
  F. Sponsorship ............................................................................................................................................... 59
  G. Comparative Efficacy and Effectiveness ................................................................................................. 59
  H. General Efficacy ....................................................................................................................................... 59
    Adalimumab .......................................................... 59
    Etanercept .............................................................. 60
    Infliximab ............................................................. 60

VI. Ulcerative colitis (UC) ............................................................................................................................... 66
  A. Summary of the Evidence .......................................................................................................................... 66
  B. Description of Studies ............................................................................................................................... 66
  C. Study Populations ....................................................................................................................................... 66
  D. Outcome Measures ..................................................................................................................................... 66
  E. Methodological Quality ............................................................................................................................. 67
  F. Sponsorship ............................................................................................................................................... 67
  G. Comparative Efficacy and Effectiveness ................................................................................................. 67
  H. General Efficacy ....................................................................................................................................... 67
    Infliximab ............................................................. 67

VII. Plaque Psoriasis ........................................................................................................................................ 70
  A. Summary of the evidence .......................................................................................................................... 70
  B. Description of Studies ............................................................................................................................... 70
  C. Study Populations ....................................................................................................................................... 70
  D. Outcome Measures ..................................................................................................................................... 70
  E. Methodological Quality ............................................................................................................................. 70
  F. Sponsorship ............................................................................................................................................... 71
  G. Comparative Efficacy and Effectiveness ................................................................................................. 71
  H. General Efficacy ....................................................................................................................................... 71
    Alefacept .............................................................. 71
    Efalizumab ............................................................ 71
    Etanercept ............................................................ 72
    Infliximab ............................................................. 72

KEY QUESTION 2: SAFETY ............................................................................................................................... 75
  A. Summary of the Evidence .......................................................................................................................... 75
  B. Overall Tolerability ..................................................................................................................................... 75
  C. Specific Adverse Events ............................................................................................................................ 77
    Serious Infections .................................................... 77
    Lymphoma and other malignancies ........................................ 79
    Congestive Heart Failure ........................................ 80
    Other Adverse Events ............................................ 81

KEY QUESTION 3: SUBGROUPS .......................................................................................................................... 88
  A. Summary of the Evidence .......................................................................................................................... 88
  B. Age ................................................................................................................................. 88
  C. Ethnicity .................................................................................................................................................. 89
  D. Sex ......................................................................................................................................................... 89
CONCLUSIONS .......................................................................................................................... 95

Gaps in the Evidence ................................................................................................................. 96

REFERENCES ................................................................................................................................ 134

TABLES
Table 1. Targeted Immune Modulators ....................................................................................... 8
Table 2. Recommended Dosage and Administration .................................................................. 10
Table 3. Criteria for the Classification of RA* (revised 1987) ....................................................... 12
Table 4. Outcome Measures and Study Eligibility Criteria ......................................................... 18
Table 5. Adjusted Indirect Comparisons of TIMs for the Treatment of RA ................... 30
Table 6. Summary of Efficacy Trials in Adult Patients with RA ................................................. 39
Table 7. Summary of Efficacy Trials in Patients with JRA ............................................................ 47
Table 8. Summary of Efficacy Trials in Adult Patients with AS .................................................. 51
Table 9. Summary of Efficacy Trials in Adult Patients with PsA ................................................ 56
Table 10. Summary of Efficacy Trials in Adult Patients with Crohn’s Disease ..................... 63
Table 11. Summary of Efficacy Trials in Adult Patients with Ulcerative Colitis ................. 69
Table 12. Summary of Efficacy Trials in Adult Patients with Plaque Psoriasis ................... 73
Table 13. Summary of Studies Assessing Adverse Events ......................................................... 83
Table 14. Summary of Studies Assessing Subgroups ................................................................. 93
Table 15. Summary of the Evidence ......................................................................................... 98

FIGURES
Figure 1. Adjusted Indirect Comparisons of Anakinra with anti-TNF Drugs for the Treatment of RA ................. 31
Figure 2. Results of literature search ....................................................................................... 101

APPENDICES
Appendix A. Search Strategy ..................................................................................................... 102
Appendix B. Studies Already Included in Meta-analyses .......................................................... 103
Appendix C. Quality Criteria .................................................................................................... 105
Appendix D. Clinical Assessment Scales Commonly Used in Targeted Immune Modulators Trials .......................................................... 107
Appendix E. Study Characteristics, Pooled Relative Risks, and Forest Plots of Meta-analyses .......................................................... 111
Appendix F. Abstract-only Studies (Not Included) ................................................................. 128
Appendix G. Characteristics of Studies with Poor Internal Validity ........................................ 132
Appendix H. Acknowledgements ........................................................................................... 133

EVIDENCE TABLES are available as an addendum to this report.

Funding:

The funding source, the Center for Evidence-based Policy, is supported by 17 organizations, including 15 state Medicaid programs. These organizations selected the topic and had input into the Key Questions for this review. The content and conclusions of this review are entirely determined by the Evidence-based Practice Center researchers. The authors of this report have no financial interest in any company that makes or distributes the products reviewed in this report.
Suggested Citation:

List of Abbreviations

ABA  Abatacept
ACR20/50/70 American College of Rheumatology, numbers refer to percentage improvement
ACR-N numeric index of the ACR response
ACT Active Ulcerative Colitis Trials
ADA adalimumab
AKA anakinra
ALE alefacept
ANA anti-nuclear antibodies
anti-ds DNA antibodies to double-stranded DNA
anti-TNF antibodies against tumor necrosis factor
AS ankylosing spondylitis
AUC area under the curve
ASA Assessment in Ankylosing Spondylitis
ASAS20 ASA 20% improvement
ASAS50 ASA 50% improvement
ASAS70 ASA 70% improvement
ASHI arthritis-specific health index
BASDAI Bath AS Disease Activity Index
BASFI Bath Ankylosing Spondylitis Functional Index
BASMI Bath Ankylosing Spondylitis Metrology Index
BSA body surface area
CAHP Childhood Arthritis Health Profile
CDAI Crohn’s Disease Activity Index
CDEIS Crohn’s Disease Endoscopy Index of Severity
CDER Center for Drug Evaluation Research
CHAQ Childhood Health Assessment Questionnaire
CHF congestive heart failure
CHQ Childhood Health Questionnaire
CI confidence interval
CRP C-reactive protein
CVD cardiovascular disease
CYP cyclosporine
DAS disease activity score
DLQI Dermatology Life Quality Index
DQOLS Dermatology Quality of Life Scales
DMARD disease-modifying antirheumatic drug
EFA Efalizumab
ESR erythrocyte sedimentation rate
ETA etanercept
EULAR European League Against Rheumatism
FDA Food and Drug Administration
HAQ Health Assessment Questionnaire
HAQ-DI Disability Index of the Health Assessment Questionnaire
HQL health-related quality of life
IBDQ Inflammatory-bowel-disease questionnaire
IgG immunoglobulin G
IgM immunoglobulin M
IL interleukin
INF infliximab
ISR injection site reaction
ITT intention to treat
JIA juvenile idiopathic arthritis
JRA juvenile rheumatoid arthritis
JCA juvenile chronic arthritis
LFT liver function test
LOCF last observation carried forward
MCS Medical Component Score
MOS Medical Outcomes Study
MTX methotrexate
NMSC non-melanoma skin cancer
N/A not applicable
NAPSI nail psoriasis and severity index
NICE National Institute for Clinical Excellence
NNT number needed to treat
NR not reported
NSAID non-steroidal anti-inflammatory drug
OR odds ratio
PASI Psoriasis Area and Severity Index
PCS Physical Component Score
PDAI Pouchitis Disease Activity Index
PGA Physician Global Assessment
PGPA Patient’s Global Psoriasis Assessment
PJRA polyarticular juvenile rheumatoid arthritis
PSA Psoriasis Symptom Assessment
PsA psoriatic arthritis
PsARC Psoriatic Arthritis Response Criteria
QALY quality-adjusted life-year
QoL quality of life
RA rheumatoid arthritis
RF rheumatoid factor
RIT rituximab
RR relative risk
SIR standard incidence ratio
s.c. subcutaneous
SF-36 Medical Outcomes Study Short Form 36 Health Survey
SJC swollen joint count
sPGA static Physician Global Assessment
TB tuberculosis
TJC tender joint count
TNF tumor necrosis factor
TNF-α tumor necrosis factor alpha
TNFβ tumor necrosis factor beta
UC ulcerative colitis
URTI upper respiratory tract infection
UTI urinary tract infection
WBC white blood cell
WESR Westergren erythrocyte sedimentation rate
Introduction

A. Targeted Immune Modulators (TIMs)
Targeted immune modulators (TIMs) – commonly referred to as biological response modifiers or simply biologics – are a relatively new category of medication used in the treatment of certain types of immunologic and inflammatory diseases, including rheumatoid arthritis (RA), juvenile rheumatoid arthritis (JRA), ankylosing spondylitis (AS), psoriatic arthritis (PsA), plaque psoriasis, Crohn’s disease, and ulcerative colitis (UC). The US Food and Drug Administration (FDA) approved the first of the biologics (infliximab) in 1998 and approved seven additional agents since that time for treating various rheumatic conditions and plaque psoriasis: etanercept (1998), anakinra (2001), adalimumab (2002), alefacept (2003), efalizumab (2003), abatacept (2005), and rituximab (2006). Table 1 summarizes currently approved biologics in the US, including trade name, manufacturer, route of administration, therapeutic mechanism of action, and approved (labeled) uses.

Table 1. Targeted Immune Modulators

<table>
<thead>
<tr>
<th>Generic Name</th>
<th>US Trade Name</th>
<th>Manufacturer</th>
<th>Route</th>
<th>Half-life</th>
<th>Onset of Action</th>
<th>Mechanism of Action</th>
<th>Labeled Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept</td>
<td>Orencia®</td>
<td>Bristol Myers Squibb</td>
<td>Intravenous</td>
<td>8-25 days</td>
<td>&gt;12 days</td>
<td>CTLA 4-Ig</td>
<td>- RA</td>
</tr>
</tbody>
</table>
| Adalimumab   | Humira®       | Abbott                | Subcutaneous    | 10-20 days| 1-14 days      | TNF inhibitor        | - RA  
|              |               |                       |                 |           |                |                     | - PsA  
|              |               |                       |                 |           |                |                     | - AS              |
| Alefacept    | Amevive®      | Biogen                | Intramuscular   | 11-12 days| 30-60 days     | CD2 antagonist       | - Plaque Psoriasis|
| Anakinra     | Kineret®      | Amgen                 | Subcutaneous    | 7-8 hours | 7-21 days      | IL-1 receptor antagonist | - RA         |
| Efalizumab   | Raptiva®      | Genentech             | Subcutaneous    | 6.2 days  | 14 days        | CD11a inhibitor      | - Plaque Psoriasis|
| Etanercept   | Enbrel®       | Amgen, Wyeth Immunex  | Subcutaneous    | 4.8 days  | 1-28 days      | TNF inhibitor        | - RA  
|              |               |                       |                 |           |                |                     | - JRA  
|              |               |                       |                 |           |                |                     | - PsA  
|              |               |                       |                 |           |                |                     | - AS              |
|              |               |                       |                 |           |                |                     | - Plaque Psoriasis   |
| Infliximab   | Remicade®     | Centocor              | Intravenous     | 9.8 days  | 2-14 days      | TNF inhibitor        | - RA  
|              |               |                       |                 |           |                |                     | - Crohn’s Disease  |
|              |               |                       |                 |           |                |                     | - PsA              |
|              |               |                       |                 |           |                |                     | - AS               |
|              |               |                       |                 |           |                |                     | - Ulcerative        |
TIMs work by selectively blocking mechanisms involved in the inflammatory and immune response. Tumor necrosis factor (TNF) inhibitors block specific proinflammatory mediators known as cytokines. Adalimumab, etanercept, and infliximab target TNF-α. Adalimumab is a fully human monoclonal antibody that binds specifically to TNF-α, blocking its interaction with both the p55 and p75 cell surface TNF receptor. Etanercept is a soluble dimeric form of the p75 TNF-α receptor linked to the Fc portion of human immunoglobulin G1 (IgG1). It exerts its action by binding circulating TNF-α and lymphotoxin-α and preventing it from interacting with a cell surface receptor. Infliximab is a chimeric (mouse/human) anti-TNF-α antibody that binds both the circulating and transmembrane forms of TNF-α, thereby preventing binding with the receptor; infliximab does not neutralize lymphotoxin alpha.

Interleukin-1 (IL-1), another naturally occurring cytokine, has both immune and pro-inflammatory actions. Anakinra is a human recombinant protein that competitively blocks the IL-1 receptor, thus blocking various inflammatory and immunological responses.

The immunosuppressant agents abatacept, alefacept, and efalizumab produce their immune response by interfering with T lymphocyte activation. Abatacept is a soluble fusion protein that consists of the extracellular domain of human cytotoxic T lymphocyte-associated antigen (CTLA-4) and the modified Fc portion of IgG1. Alefacept is a dimeric fusion protein that consists of the extracellular CD2-binding portion of the human leukocyte function antigen (LFA-3) and the Fc portion of human IgG1. Efalizumab is a recombinant humanized IgG1 monoclonal antibody that binds to human CD11a and inhibits the binding of LFA-1 to intercellular adhesion molecule-1 (ICAM-1).

Rituximab, a chimeric murine/human monoclonal antibody, works by binding to the CD20 antigen found on the surface of B lymphocytes. B-cells are believed to play a role in autoimmune and inflammatory processes, such as those involved in RA. As this report was going to press, the FDA issued an alert highlighting important emerging safety information on rituximab for healthcare professionals. Two patients have died after being treated with rituximab for systemic lupus erythematosus (SLE). The cause of death was a viral infection of the brain called progressive multifocal leukoencephalopathy.
Although the treatment of SLE constitutes an off-label use of rituximab, the risk of reactivation and exacerbation of viral infections is likely generalizable to patients with other conditions, given the immunosuppressive nature of rituximab. The FDA is working with the manufacturer of rituximab to gather additional information and will update the current analysis.

Because they have a similar mechanism of action, adalimumab, etanercept, and infliximab are used interchangeably in the treatment of RA, although the clinical response to the different agents can vary widely in an individual patient. Alefacept, anakinra, and efalizumab each produces its effect by affecting a different point in the inflammatory and immune response cascade. Table 2 summarizes dosages and administration for different indications.

**Table 2. Recommended Dosage and Administration**

<table>
<thead>
<tr>
<th>Generic Name</th>
<th>Indication</th>
<th>Dosage and Administration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abatacept</td>
<td>RA</td>
<td>Intravenous infusion dosed according to body weight (&lt; 60kg = 500mg; 60-100kg = 750mg; &gt; 100kg = 1000mg); dose repeated at 2 weeks and 4 weeks after initial dose, and every 4 weeks thereafter</td>
</tr>
<tr>
<td>Adalimumab</td>
<td>RA</td>
<td>40 mg every other week as subcutaneous injection; may increase to 40 mg per week for adalimumab monotherapy</td>
</tr>
<tr>
<td></td>
<td>PsA</td>
<td>40 mg every other week as subcutaneous injection</td>
</tr>
<tr>
<td>Alefacept</td>
<td>Plaque Psoriasis</td>
<td>15 mg given once weekly as an intramuscular injection. Treatment should be continued for 12 weeks; re-treatment with an additional 12 week course may be initiated provided that CD4+ T lymphocytes counts are &gt; 250 cells/µL and a 12-week interval has passed since the end of the initial treatment cycle</td>
</tr>
<tr>
<td>Anakinra</td>
<td>RA</td>
<td>100 mg daily as subcutaneous injection; dose should be decreased to 100 mg every other day in renal insufficiency</td>
</tr>
<tr>
<td>Efalizumab</td>
<td>Plaque Psoriasis</td>
<td>Initial 0.7 mg/kg subcutaneous injection followed by weekly doses of 1 mg/kg (not to exceed total of 200 mg)</td>
</tr>
<tr>
<td>Etanercept</td>
<td>RA</td>
<td>25 mg twice weekly as subcutaneous injections or 50 once weekly as subcutaneous injection</td>
</tr>
<tr>
<td></td>
<td>PsA</td>
<td>0.8 mg/kg per week (maximum 50 mg per week) given as one or two subcutaneous injections</td>
</tr>
<tr>
<td></td>
<td>AS</td>
<td>Plaque Psoriasis is 50 mg given twice weekly (administered 3 or 4 days apart) as a subcutaneous injection for 3 months, followed by 50 mg weekly</td>
</tr>
<tr>
<td></td>
<td>JRA (patients 4-17 years)</td>
<td>0.8 mg/kg per week (maximum 50 mg per week) given as one or two subcutaneous injections</td>
</tr>
<tr>
<td>Infliximab</td>
<td>RA</td>
<td>Adult: 3 mg/kg intravenous 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 Pediatric (6-17 years): 5 mg/kg intravenous infusion at 0, 2, and 6 weeks followed by maintenance every 8 weeks</td>
</tr>
<tr>
<td></td>
<td>Crohn’s Disease</td>
<td>5 mg/kg intravenous infusion at 0, 2, and 6 weeks followed by maintenance every 8 weeks thereafter; may increase to 10 mg/kg</td>
</tr>
<tr>
<td></td>
<td>PsA</td>
<td>5 mg/kg intravenous infusion at 0, 2, and 6 weeks followed by maintenance every 8 weeks thereafter</td>
</tr>
<tr>
<td>Generic Name</td>
<td>Indication</td>
<td>Dosage and Administration</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>AS</td>
<td>AS</td>
<td>5 mg/kg intravenous infusion at 0, 2, and 6 weeks followed by maintenance every 6 weeks thereafter</td>
</tr>
<tr>
<td>Active UC</td>
<td>Active UC</td>
<td>5 mg/kg induction regimen at 0, 2, and 6 weeks followed by a maintenance regimen of 5 mg/kg every 8 weeks thereafter</td>
</tr>
<tr>
<td>Plaque Psoriasis</td>
<td>Plaque Psoriasis</td>
<td>5 mg/kg induction regimen at 0, 2, and 6 weeks followed by a maintenance regimen of 5 mg/kg every 8 weeks thereafter</td>
</tr>
<tr>
<td>Rituximab</td>
<td>RA</td>
<td>1000 mg intravenous infusion on days 1 and 15 in combination with methotrexate</td>
</tr>
</tbody>
</table>

In this report, we review the comparative effectiveness, safety, and tolerability of TIMs. Our review covers the use of these drugs in adult patients with RA, AS, PsA, Crohn’s disease, UC, plaque psoriasis, and pediatric patients with JRA. The next section briefly describes the epidemiology and pathophysiology of these conditions, as well as clinical features, assessment methods, management goals, and treatment strategies. Furthermore, we review the role of the targeted immune modulators in treating patients with these diseases.

**B. Rheumatoid Arthritis (RA)**

RA is an autoimmune disease that affects about one percent of the population worldwide. The exact etiology of RA is not completely understood, but genetic susceptibility factors have been described in certain populations. The hallmarks of the disease are inflammation of the synovial tissues with progressive erosion of bone leading to malalignment of the joint and disability in most cases. Studies have shown the importance of CD4+ T cells, B cells, and cytokines in the pathogenesis of RA. TNF-α plays a central role in the pathobiology of RA. It is an important regulator of other pro-inflammatory molecules and stimulates the secretion of matrix metalloproteinases. It also exerts a direct effect on the multiple tissues inside the joint including chondrocytes, macrophages, synovial fibroblasts, and osteoclasts. Together, its action leads to inflammation and the formation of pannus, a localized mass of tissue that causes localized joint destruction.¹

The diagnosis of RA is primarily a clinical one. Constitutional symptoms, such as fatigue and low grade fevers, are common before the onset of joint swelling and pain. Joint stiffness is almost always present and is frequently most severe after periods of prolonged rest. The disease tends to affect the small joints of the hands and feet first in a symmetric pattern, but other joint patterns are often seen. In a subset of patients, RA can be a devastating disease with numerous extra-articular manifestations. Severe disease may be complicated by involvement of the eyes, lungs, nerves, and the cardiovascular system.
A serum rheumatoid factor is present in up to 75 percent of patients with RA but is frequently negative in early disease. A more specific marker, anti-cyclic citrullinated peptide (CCP) antibody, has recently been described and may be a useful marker in patients with early disease.\(^2\) Table 3 presents the classification criteria for RA proposed by the American College of Rheumatology (ACR). These criteria were developed for use in clinical trials, but may be relatively insensitive in early disease.

<table>
<thead>
<tr>
<th>Table 3. Criteria for the Classification of RA* (revised 1987)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Morning stiffness lasting greater than one hour</td>
</tr>
<tr>
<td>2. Arthritis in 3 or more joint areas</td>
</tr>
<tr>
<td>3. Arthritis of the hand joints (metacarpophalangeal [MCP], proximal interphalangeal [PIP], wrists)</td>
</tr>
<tr>
<td>4. Symmetric arthritis</td>
</tr>
<tr>
<td>5. Rheumatoid nodules</td>
</tr>
<tr>
<td>6. Serum rheumatoid factor</td>
</tr>
<tr>
<td>7. Radiographic changes: erosions or unequivocal periarticular osteopenia</td>
</tr>
</tbody>
</table>

*Patients are said to have RA if they meet 4 of 7 criteria.\(^3\)

Treatment is aimed at controlling pain and inflammation and ultimately, slowing or arresting the progression of joint destruction. The key to successful management of RA is the early identification of the disease and the rapid institution of effective therapies.\(^4\) Methotrexate (MTX) is the cornerstone of most RA treatment regimens as it has demonstrated good disease control and tolerability. However, MTX toxicity may limit the use of MTX, and many patients do not adequately respond to MTX monotherapy. In patients with persistent disease despite aggressive management with oral agents, biologic agents, often in combination with MTX, are now considered the standard of care. Lifelong therapy is usually necessary.

**C. Juvenile Rheumatoid Arthritis (JRA)**

JRA is a form of arthritis that, by definition, lasts at least 6 weeks in a child under the age of 16. It is a systemic disease with a variable presentation and has three established subtypes: pauciarticular (<5 joints involved), polyarticular (\(\geq\) 5 joints involved), and systemic (arthritis with fever and a rash).\(^5\)

Joint pain, stiffness, and swelling are the hallmarks of JRA. Children with systemic disease often present with constitutional symptoms such as fever or rash. Similar findings may be seen in polyarticular disease but are rare with pauciarticular presentation. Uveitis, an inflammatory disease of the eye, is common. Children with the most severe forms of JRA may have significant disability from progressive destructive arthritis. Long-term consequences of the disease include growth disturbances, deformity of the joints, and blindness.
Initial therapeutic strategies are aimed at decreasing pain and swelling and improving the child’s functional status. Non-steroidal anti-inflammatory drugs (NSAIDs) are first line therapy and are usually fairly well tolerated in children. Systemic steroids are usually avoided, if possible, because of adverse effects on bone growth. However, intra-articular steroid injections can be an effective strategy, particularly if only a few joints are afflicted with active disease. As in RA, oral disease-modifying antirheumatic drugs (DMARDs) are used next, with MTX being the most widely used. When the disease is resistant to oral therapies, biologic agents are indicated.

**D. Ankylosing Spondylitis (AS)**

AS is a chronic inflammatory arthritis with primary involvement of the axial skeleton and prominent involvement of the spine and sacroiliac joints. Peripheral joint disease can occur and may be destructive in some cases. The peak age of onset is in the 20s, and men are affected more frequently than women by a ratio of about 3 to 1. The onset is indolent with prominent stiffness in the low back, which is characteristically worse at night and in the early morning. The sacroiliac joints are usually the first joints involved, and the disease is characterized by progressive involvement of the spine. Enthesitis, inflammation of the insertion of ligaments and tendons on bones, is one of the hallmarks of the disease.

Existing diagnostic criteria are relatively insensitive and have limited utility in clinical practice. AS usually presents with inflammatory back pain and stiffness in a young adult, although 20 percent present with peripheral joint involvement and more than 50 percent have joints other than the spine affected at some stage. Radiographs of the sacroiliac joints, when abnormal, can be useful in assessing the presence of AS; however, they are frequently normal in early disease. Over time, patients with AS develop progressive fusion of the spine with resultant deformity and disability.

For years NSAIDs were the standard of care for the treatment of AS, as they are effective in treating pain and stiffness. However, they do not have any effect on disease progression. Traditional DMARDs have been used, mostly because a lack of other more effective therapies, although they are usually ineffective in treating spinal arthritis. As TNF has been implicated in the pathophysiology of AS, biologic agents targeting TNF have become a standard treatment approach. Studies are under way to assess whether treatment with these agents affects the natural history of AS.
**E. Psoriatic Arthritis (PsA)**
PsA is a chronic inflammatory arthritis associated with the skin disease psoriasis. In most cases, the psoriasis predates the onset of the PsA. The presentation, however, is highly variable. In all cases, symptoms include pain and stiffness in the affected joint as well as joint line tenderness, swelling, and sometimes loss of range of motion. Pitting of the fingernails often correlates with the extent and severity of the disease. Dactylitis, swelling of a whole digit, is a characteristic clinical finding. Enthesitis, spondylitis, sacroiliitis, and inflammatory eye disease (iritis, uveitis) may occur.

The etiology and pathogenesis of psoriasis and PsA are not completely understood, but genetic, immunologic, and environmental factors are all likely to play a role. The first line of treatment is NSAIDs, although in most cases DMARDs are necessary. Corticosteroids may be used but do not have much of a role in chronic disease management in psoriatic disease. If disease continues to be active despite the use of NSAIDS, MTX, or other oral DMARDS, biologics may be indicated.

**F. Crohn’s Disease**
Crohn’s disease is a condition of the bowel causing inflammation involving the full thickness of the bowel wall. This may occur at any point from the mouth to the anus. This chronic inflammation leads to fibrosis and obstructive symptoms with sinus tracts and fistulae. Fistulizing disease is a serious complication of Crohn’s disease; it is basically abnormal communication between the gut and the skin or other internal organs, with small bowel or colonic contents draining to the skin or other organs. Abdominal pain and diarrhea, with or without bleeding, are characteristic of the disease. Constitutional symptoms are very common, predominantly fatigue and weight loss. Nonspecific digestive symptoms may predate the onset of clinically overt disease. Extra-intestinal symptoms may occur and include inflammatory eye disease, arthritis, and sclerosing cholangitis. Clinical diagnosis is made on the basis of history and physical examination and is confirmed on endoscopy and biopsy of the involved segment of the GI tract. Patients with aggressive or poorly controlled disease may suffer numerous complications; these include severe hemorrhage, intestinal obstruction, perforation, development of fistulae and abscess formation, malabsorption with nutritional deficiencies, and rarely, malignancy.

Treatment is aimed at controlling the inflammation and preventing complications. Mild disease may be controlled with 5-aminosalicylate (ASA) drugs or antibiotics. If the disease is resistant to these interventions or is more severe, corticosteroids are frequently used. If symptoms persist despite steroids or if the disease flares on tapering the steroids, immunomodulatory agents (azathioprine, 6-
mercaptopurine [6-MP], and MTX) often are instituted. Biologics may be warranted in patients with moderate to severe active Crohn’s disease who have had inadequate response to conventional therapy. It is recommended that medical therapy be exhausted before surgical therapy is considered, except in cases of catastrophic complications such as acute colonic obstruction, massive hemorrhage, or bowel perforation.

**G. Ulcerative Colitis (UC)**

UC is a chronic inflammatory bowel disease (IBD), that is characterized by mucosal ulceration, rectal bleeding, diarrhea, and abdominal pain limited to the colon and rectal areas, unlike Crohn’s disease which causes inflammation deeper within the intestinal wall and can occur in other parts of the digestive system including the small intestine, mouth, esophagus, and stomach. The most common symptoms of ulcerative colitis are abdominal pain and bloody diarrhea. Clinical diagnosis is most accurately made with colonoscopy or sigmoidoscopy.

Treatment is aimed at reducing and maintaining remission of symptoms and inflammation. Mild disease may be controlled with oral and/or topical 5-aminosalicylate (ASA) drugs. If the disease is resistant to these interventions or is more severe, corticosteroids are frequently used. In addition, infliximab has been approved by the FDA for treatment of moderate to severe ulcerative colitis. Indications for surgery include excessive bleeding, perforation, carcinoma and toxic colitis. About 25 to 40 percent of ulcerative colitis patients must eventually have their colons removed.

**H. Plaque Psoriasis**

Plaque psoriasis is a chronically recurring, debilitating inflammatory disease that affects the skin, scalp, and joints. It is characterized by erythrosquamous skin lesions and ranges in severity from mild to severe. Patients with moderate to severe disease experience significant deterioration of quality of life. The exact pathogenesis of plaque psoriasis is still unknown; however, pathophysiological evidence suggests that an overproduction of proinflammatory cytokines plays an important role. In particular, TNF levels are increased in psoriatic lesions compared with healthy skin.

The severity of plaque psoriasis is most commonly classified based on the percentage of body surface area (BSA) involved. Severe psoriasis is generally defined as more than 10 percent BSA affected.
The goal of plaque psoriasis treatment is to gain control of the disease process, decrease the percentage of body surface involved, and achieve and maintain long-term remission. Conventional therapy includes topical treatments (e.g. topical corticosteroids, calcipotriene, tazarotene), phototherapy (e.g. broadband UVB [ultraviolet B light], narrowband UVB, PUVA [psoralen plus ultraviolet A light]), and systemic therapy (e.g., MTX, cyclosporine, acitretine). In addition, biologic agents such as alefacept, efalizumab, etanercept, and infliximab have been approved by the FDA for the treatment of moderate to severe plaque psoriasis.

I. Scope and Key Questions
The purpose of this review is to help policy makers and clinicians make informed choices about the use of targeted immune modulators. We compare the efficacy, effectiveness, and safety (adverse events) of abatacept, adalimumab, alefacept, anakinra, efalizumab, etanercept, infliximab, and rituximab in patients with RA, JRA, AS, PsA, Crohn’s disease, UC, and plaque psoriasis.

The participating organizations of the Drug Effectiveness Review Project (DERP) are responsible for ensuring that the scope of the review reflects the populations, drugs, and outcome measures of interest to their constituencies. The Oregon Evidence-based Practice Center initially prepared preliminary key questions identifying the populations, interventions, and outcomes of interest, and we based the eligibility criteria for studies on these preliminary questions. Representatives of organizations participating in the DERP, in conjunction with experts in the fields of health policy, rheumatology, pharmacotherapy, and research methods reviewed, revised, and approved the questions and outcome measures. The participating organizations approved the following key questions:

1. How do included drugs compare in their effectiveness for alleviating symptoms and stabilizing the disease in patients with RA, JRA, AS, PsA, Crohn’s disease, UC, and plaque psoriasis?

2. What are the comparative incidence and severity of complications of these drugs?

3. Do the included drugs differ in effectiveness or adverse events in different age, sex, or ethnic groups, or in patients taking other commonly prescribed drugs?
The first key question addresses the issue of effectiveness: do the biologics differ in their effects under real-life circumstances? This report addresses both efficacy (i.e., whether biologics differ in their effects under ideal or highly controlled circumstances) and effectiveness. We distinguish between *efficacy* (*explanatory*) studies and *effectiveness* (*pragmatic*) studies; studies conducted in community-based settings that use less stringent eligibility criteria (i.e., broad range of population characteristics and disease severity), have long follow-up periods (i.e., greater than one year), and assess health outcomes are characterized as *effectiveness* studies. Studies conducted in more highly selected populations over shorter periods of time are characterized as *efficacy* studies. We summarize the results of efficacy and effectiveness studies separately as the results of effectiveness studies are more generalizable than results from highly selected populations (i.e., efficacy studies). However, effectiveness studies may have lower internal validity because of a higher risk of bias.

For assessing efficacy, effectiveness, and safety our review includes methodologically valid controlled clinical trials, placebo-controlled trials, fair- or good-quality systematic reviews, and fair- or good-quality observational studies. Table 4 summarizes outcome measures and study eligibility criteria.
### Table 4. Outcome Measures and Study Eligibility Criteria

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Outcome Measures</th>
<th>Study Eligibility Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Efficacy / Effectiveness</strong></td>
<td>Health outcomes:</td>
<td>• Outpatient study population</td>
</tr>
<tr>
<td></td>
<td>• Quality of Life</td>
<td>• Head-to-head randomized controlled clinical trials or meta-analyses comparing one TIM to another</td>
</tr>
<tr>
<td></td>
<td>• Functional capacity</td>
<td>o Good or fair quality</td>
</tr>
<tr>
<td></td>
<td>• Pain</td>
<td>o ≥ 3 months study duration</td>
</tr>
<tr>
<td></td>
<td>• Reduction in the number of swollen or tender joints</td>
<td>o N ≥ 100</td>
</tr>
<tr>
<td></td>
<td>• Response</td>
<td>• When sufficient evidence was not available for head-to-head comparisons we evaluated placebo-controlled trials</td>
</tr>
<tr>
<td></td>
<td>• Remission</td>
<td>o Good or fair quality</td>
</tr>
<tr>
<td></td>
<td>• Reduction of affected body surface area</td>
<td>o ≥ 3 months study duration</td>
</tr>
<tr>
<td></td>
<td>• Hospitalizations</td>
<td>o N ≥ 100</td>
</tr>
<tr>
<td></td>
<td>• Mortality</td>
<td>• Controlled observational studies were reviewed for quality of life, functional capacity, hospitalizations and mortality - outcome measures rarely assessed in controlled trials</td>
</tr>
<tr>
<td></td>
<td>If no studies with health outcomes were available, we included intermediate outcomes:</td>
<td>o Good or fair quality</td>
</tr>
<tr>
<td></td>
<td>• Radiological outcomes</td>
<td>o ≥ 12 months study duration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>o N ≥ 100</td>
</tr>
<tr>
<td><strong>Safety/ Tolerability</strong></td>
<td>• Overall adverse events</td>
<td>• Head-to-head randomized controlled clinical trials or meta-analyses comparing one TIM drug to another</td>
</tr>
<tr>
<td></td>
<td>• Withdrawals because of adverse events</td>
<td>o Good or fair quality</td>
</tr>
<tr>
<td></td>
<td>• Serious adverse events</td>
<td>o ≥ 3 months study duration</td>
</tr>
<tr>
<td></td>
<td>• Specific adverse events, including:</td>
<td>o N ≥ 100</td>
</tr>
<tr>
<td></td>
<td>- serious infectious diseases</td>
<td>• Placebo-controlled trials</td>
</tr>
<tr>
<td></td>
<td>- lymphoma</td>
<td>o Good or fair quality</td>
</tr>
<tr>
<td></td>
<td>- CHF</td>
<td>o ≥ 3 months study duration</td>
</tr>
<tr>
<td></td>
<td>- autoimmunity</td>
<td>o N ≥ 100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Observational studies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>o Good or fair quality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>o ≥ 6 months study duration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>o N ≥ 100</td>
</tr>
</tbody>
</table>

CHF: congestive heart failure; TIM: targeted immune modulator
As equipotency among the reviewed biologics is not well established, we assume that comparisons made within the recommended dosing range are appropriate (Table 2). Dose comparisons made outside the recommended daily dosing range are acknowledged in our report, but we do not use them to determine the quality of the evidence.

Under normal circumstances, intravenous TIMs are rarely administered in primary care practices. They are used by specialists such as rheumatologists, gastroenterologists, and sometimes dermatologists. Some agents may be patient-administered with proper training.
METHODS

A. Literature Search
To identify articles relevant to each key question we searched MEDLINE, Embase, The Cochrane Library, and the International Pharmaceutical Abstracts; we used either Medical Subject Headings (MeSH or MH) as search terms when available or key words when appropriate. We combined terms for selected indications (RA, JRA, AS, PsA, Crohn’s disease, UC, plaque psoriasis), drug interactions, and adverse events with a list of eight specific TIMs (abatacept, adalimumab, alefacept, anakinra, efalizumab, etanercept, infliximab, rituximab). We limited the electronic searches to “human” and “English language”; we searched sources from 1980 to 2006 (August) to delimit literature relevant to the scope of our topic.

We used the National Library of Medicine publication type tags to identify reviews, randomized controlled trials (RCTs), and meta-analyses; we also manually searched reference lists of pertinent review articles and letters to the editor. All citations were imported into an electronic database (EndNote, version 8.0). Additionally, we hand-searched the Center for Drug Evaluation and Research (CDER) database to identify unpublished research submitted to the FDA.

Further, the Center for Evidence-based Policy at the Oregon Health and Science University (OHSU) contacted pharmaceutical manufacturers and invited them to submit dossiers, including citations, using a protocol available at www.ohsu.edu/drugeffectiveness. We received dossiers from five pharmaceutical companies (Abbott Laboratories, Amgen Pharmaceuticals, Bristol Myers Squibb, Centocor, Genentech, Wyeth/Amgen Pharmaceuticals).

Our searches found 2,711 citations, unduplicated across databases; we found an additional 118 articles from manually reviewing the reference lists of pertinent review articles and an additional 14 articles in the pharmaceutical dossiers. We found six new citations from reviewing public comments. The total number of citations included in the database was 2,849. For further details on the search strategy, see Appendix A.
B. Study Selection

Two people independently reviewed abstracts; if both reviewers agreed that the study did not meet eligibility criteria, it was excluded. We obtained the full text of all remaining articles. Records were considered for exclusion if they did not meet pre-established eligibility criteria with respect to study design or duration, patient population, interventions, outcomes, and comparisons to medications outside our scope of interest.

With respect to study design we took a “best evidence” approach for this review. Results from well-conducted, head-to-head trials provide the strongest evidence to compare drugs with respect to effectiveness, efficacy, and adverse events; head-to-head trials were defined as those comparing one TIM with another. RCTs of at least 3 months’ duration having an outpatient study population with a total sample size greater than 100 participants were eligible for inclusion.

If we could not find sufficient evidence of efficacy or effectiveness from at least one randomized, double-blinded trial for a certain indication, we reviewed other study designs as needed. Thus, to present the best available evidence, we also reviewed experimental studies with fewer than 100 participants or with an open-label design. In addition, we reviewed large (n > 100), well-conducted, observational studies (cohort studies, case control studies, case series) with a follow-up of at least 1 year to augment findings from experimental studies. Long-term observational studies can provide evidence on outcomes that may be difficult to observe in RCTs due to limitations in sample sizes and study durations. Furthermore, observational data can provide information whether treatment effects observed in RCTs can be translated to less selected populations. Nevertheless, the strength of evidence of these results for comparing different drugs must be rated lower than results from the most preferred type of trial.

If no head-to-head evidence was published, we reviewed placebo-controlled trials for indications of interest. We reviewed all placebo-controlled trials to provide an overview of efficacy without taking drug equivalency into account. We compared results of approved dosing ranges, but no evidence on exact comparative dosing is currently available. Study populations, disease severity, and concomitant treatments can differ considerably across placebo-controlled trials. Comparisons of treatment effects across trials must, therefore, be made with caution.

We included meta-analyses in the evidence report if they were relevant to a key question and of good or fair methodological quality (based on the QUORUM statement). We did not summarize individual studies in evidence tables if they were included in a high-quality meta-analysis. We excluded meta-
analyses that were not based on a comprehensive systematic literature search or did not maintain the units of the studies in their statistical analyses. We checked our database to guarantee that our literature search had detected trials included in any meta-analyses that we discarded and obtained any missing articles.

For adverse events we included both experimental and observational studies. For observational studies we included those with large sample sizes (≥ 100 patients) that lasted at least 6 months and reported an included outcome.

We initially reviewed studies with health outcomes as the primary outcome measures. Outcomes were quality of life, functional capacity, alleviation of symptoms, hospitalizations, and mortality. If no study measuring health outcomes was available for a particular indication or population subgroup, we included intermediate outcomes (e.g., radiological changes). Safety outcomes included overall and specific adverse events (e.g., serious infections, lymphoma, autoimmunity), withdrawals attributable to adverse events or lack of efficacy, and drug interactions.

We included a total of 504 articles on an abstract level and retrieved those as full text articles for background information or to be reviewed for inclusion into the evidence report. We did not summarize studies that were included in a high-quality meta-analysis (listed in Appendix B).

C. Data Abstraction
We designed and used a structured data abstraction form to ensure consistency in appraisal for each study. Trained reviewers abstracted data from each study and assigned an initial quality rating. A senior reviewer read each abstracted article, evaluated the completeness of the data abstraction, and confirmed the quality rating. We abstracted the following data from included trials: study design, eligibility criteria, intervention (drugs, dose, duration), additional medications allowed, methods of outcome assessment, population characteristics, sample size, loss to follow-up, withdrawals attributed to adverse events, results, and adverse events reported. We recorded intention-to-treat results if available.

D. Quality Assessment
We assessed the internal validity (quality) of trials based on predefined criteria (Appendix C) developed by the US Preventive Services Task Force (ratings: good-fair-poor) and the National Health Service Centre for Reviews and Dissemination. External validity (generalizability) was assessed and reported
but did not influence quality ratings. We did not rate the quality of descriptive studies (case series, database reviews).

Two independent reviewers assigned quality ratings; they resolved any disagreements by discussion and consensus or by consulting a third, independent party. Elements of internal validity assessment included, among others, randomization and allocation concealment, similarity of compared groups at baseline, use of intention-to-treat analysis, and overall and differential loss to follow-up.

Loss to follow-up was defined as the number of persons randomized who did not reach the endpoint of the study, independent of the reason and the use of intention-to-treat analysis. We adopted no formal cut-off point of loss to follow-up since many studies defined withdrawals due to acute worsening of the disease as an outcome measure. Trials that had a fatal flaw in one or more categories were rated poor quality and not included in the analysis of the evidence report; trials that met all criteria were rated good quality. The majority of trials received a quality rating of fair. This includes studies that presumably fulfilled all quality criteria but did not report their methodologies to an extent that answered all of our questions. Therefore, the “fair quality” category includes trials with quite different strengths and weaknesses and a range of validity.

**E. Data Synthesis**

Throughout this report we synthesized the literature qualitatively. If data were sufficient, we augmented findings with quantitative analyses. We conducted meta-analyses of data for placebo-controlled trials that were fairly homogenous in study populations and outcome assessments. Our outcome measure of choice for RA was the relative risk (RR) of achieving an ACR 20/50/70 response (American College of Rheumatology [ACR], numbers refer to percentage improvement [see Appendix D for a summary of different scales]). We did not find sufficient data to pool results of the Health Assessment Questionnaire (HAQ) or other measures of health-related quality of life. We chose the ACR 50 outcome measure because response to treatment can be viewed as a close proxy to health outcomes. Therefore, such an outcome measure has more clinical significance than a comparison of mean changes of scores on rating scales. A 50 percent improvement on the ACR scale (i.e., an ACR50 response) is commonly viewed as a clinically significant response.
For each meta-analysis, we conducted a test of heterogeneity ($I^2$ statistic) and applied both a random and a fixed effects model. We report the random effects model results if moderate or high heterogeneity ($I^2 > 30\%$) was present. In addition, we calculated the number needed to treat (NNT) based on the pooled risk difference.

We assessed publication bias using funnel plots and Kendell’s tests. However, given the small number of component studies in our meta-analyses, results of these tests must be viewed cautiously. All statistical analyses were conducted using StatsDirect, version 2.3.8.

Because only limited head-to-head evidence on TIMs was available, we conducted adjusted indirect comparisons when data was sufficient and trials were of similar design, conducted in similar settings with a comparable patient population. We based these analyses on the method proposed by Bucher et al.\textsuperscript{21} Evidence suggests that adjusted indirect comparisons agree with head-to-head trials if component studies are similar and treatment effects are expected to be consistent in patients included in different trials.\textsuperscript{22, 23} Nevertheless, findings must be interpreted cautiously.
RESULTS

We identified 2,849 citations from searches and reviews of reference lists. In total we included 112 studies: 57 RCTs, four observational extensions of RCTs, five meta-analyses, 43 observational studies, and four studies of other design (pooled data analysis). Furthermore, we retrieved 165 articles for background information.

Reasons for exclusions were based on eligibility or methodological criteria (Figure 1, QUORUM Tree). We excluded three studies that originally met eligibility criteria but were later rated as poor quality for internal validity (Appendix G).

Of the 112 included studies, 57 percent were financially supported by pharmaceutical companies, 11 percent were funded by governmental agencies or independent funds, and 13 percent received both pharmaceutical and government funding. We could not determine a funding source for 19 percent of the included studies.

KEY QUESTION 1
How do included drugs compare in their effectiveness for alleviating symptoms and stabilizing the disease in patients with rheumatoid arthritis, juvenile rheumatoid arthritis, ankylosing spondylitis, psoriatic arthritis, Crohn’s disease, and plaque psoriasis?

We included 57 RCTs, one non-randomized trial, four meta-analyses, and six observational studies. No RCTs were head-to-head trials. One study was characterized as an effectiveness trial.24 Most of the included efficacy studies were conducted in narrowly defined populations and/or were limited to less than 1 year of follow-up.

I. Rheumatoid Arthritis (RA)
The following drugs are currently approved by the FDA for the treatment of RA: abatacept, adalimumab, anakinra, etanercept, infliximab, and rituximab.
A. Summary of the Evidence
Overall, the evidence on the comparative effectiveness of TIMs for the treatment of RA is fair to poor. We found only one head-to-head study, which was a non-randomized, open-label effectiveness trial comparing etanercept to infliximab. Etanercept had significantly greater ACR 20 response rates at 3 and 6 months than infliximab, however, no differences existed after 1 year. Two large observational studies also reported numerically greater response rates for etanercept than for infliximab after up to 5 years of follow-up. Otherwise, no evidence directly comparing the efficacy and safety of one TIM to another could be found. Adjusted indirect comparisons of randomized placebo-controlled trials suggest that no substantial differences exist among the efficacy of adalimumab, etanercept, and infliximab. Point estimates favor adalimumab, etanercept, and infliximab over anakinra. However, differences do not reach statistical significance in adjusted indirect comparisons which is likely attributable to a lack of power. Adjusted indirect comparisons of anti-TNF drugs as a class compared to anakinra result in a statistically significantly greater efficacy of anti-TNF drugs on ACR 20 but not on ACR 50. These findings are largely consistent with a meta-analysis and adjusted indirect comparisons conducted by the UK Health Technology Assessment Programme.

Good to fair evidence exists from meta-analyses and large RCTs that abatacept, adalimumab, anakinra, etanercept, infliximab, and rituximab are significantly more efficacious than placebo for the treatment of RA. Treatment effects are large and consistent across studies. We did not find any evidence on the efficacy and safety of alefacept and efalizumab for the treatment of RA.

Although adalimumab and etanercept monotherapies failed to show a benefit relative to MTX monotherapy with respect to health outcomes (SF-36 [Medical Outcomes Study Short Form 36 Health Survey], HAQ, ASHI [Arthritis-Specific Health Index]) and ACR response rates after 52 weeks of treatment, radiographic outcomes were significantly better in TIM- than in MTX-treated patients. Two of these studies were conducted in patients with early RA.

No synergistic effects of a combination treatment of etanercept, anakinra, and MTX compared to an etanercept-MTX regimen could be detected. Furthermore, the frequency of serious adverse events was substantially higher in the etanercept-anakinra combination groups. However, this finding is based on one trial.
B. Description of Studies
For RA, we did not find any head-to-head RCTs comparing one TIM to another. We found one non-randomized, open-label trial that assessed the long-term effectiveness and safety of etanercept, infliximab, and leflunomide. This study could be characterized as an effectiveness trial. In addition, we included two prospective cohort studies and one retrospective cohort study on the comparative effectiveness of etanercept and infliximab. Furthermore, we included four meta-analyses of placebo-controlled trials, and twenty RCTs that were not included in any meta-analysis. We did not find any studies on alefacept and efalizumab. Included studies are presented in Table 6.

C. Study Populations
All patients suffered from active RA. However, the definition of active disease varied across studies. The non-randomized study and the prospective cohort studies were population-based and enrolled patients who had a diagnosis of RA based on the clinical judgment of the treating physician and who had failed to respond to at least one DMARD. Most RCTs employed the ACR criteria to classify the diagnosis of RA. Some trials, however, used stricter eligibility criteria. Disease duration and concomitant treatments also varied across studies. Most patients used NSAIDS or oral corticosteroids in addition to the study medication. The majority of trials enrolled patients who had failed at least one DMARD treatment or were on a stable dose of MTX with unsatisfactory response. Three studies examined the efficacy of TIMs in patients with early RA and no prior MTX exposure. One RCT evaluated the efficacy and safety of a combination treatment of etanercept and anakinra. Patients with an autoimmune disease other than RA, a history of active listeriosis or mycobacterial infection, or recent antibiotic treatment were generally excluded from studies.

D. Outcome Measures
All trials assessed response rates as defined by the ACR or by the European League Against Rheumatism (EULAR). These scales (ACR20/50/70, DAS28 [Disease Activity Score]) combine measures of global disease activity with counts of tender and swollen joints and acute phase laboratory parameters (see Appendix D). In addition, most studies evaluated health outcomes such as quality of life, functional capacity (e.g., SF-36, HAQ, ASHI), or discontinuation rates due to disease worsening. Some studies used the modified Sharp Method (radiographs of hands, wrists, and feet) to assess disease progression.
E. Methodological Quality
Study quality varied across studies. Some “fair” ratings are probably more attributable to inadequate reporting than to methodological flaws. Randomization methods and blinding were generally adequate; all studies used a double-dummy design (i.e., using an identical container for active treatment and placebo) to guarantee blinding; method of allocation concealment was rarely reported. The non-randomized trial was open-label and did not blind outcome assessors.

F. Sponsorship
All studies, except the non-randomized trial, the four meta-analyses, and two cohort studies were funded by the pharmaceutical industry.

G. Comparative Efficacy and Effectiveness
We did not identify any head-to-head RCTs. A fair, non-randomized, open-label trial assessed the efficacy and safety of etanercept (n = 166), infliximab (n = 135), and leflunomide (n = 103). This Swedish study was population-based and had minimal exclusion criteria. Study duration was 12 months. Etanercept had significantly greater ACR20 response rates at 3 months (P < 0.02) and 6 months (P < 0.05), and greater ACR50 response rates at 6 months (P < 0.005) than infliximab. No significant difference could be detected thereafter. Both, etanercept and infliximab had significantly greater response rates than leflunomide. Although patient characteristics were similar at baseline, results must be interpreted cautiously because of an increased risk of bias in such a study design. Nevertheless, findings from two large prospective cohort studies reported consistent results. Etanercept led to numerically greater response rates than infliximab after up to 3 years of follow-up.

The largest of these studies was a prospective cohort study based on the RADIUS (Rheumatoid Arthritis DMARD Intervention and Utilization Study) program. This multicenter (509 rheumatology practices in the US) registry enrolled patients who required changes in their RA treatment regimens. Data on 3,034 patients on etanercept and 660 patients on infliximab were available for analysis after 12 months of follow-up. Etanercept-treated patients had numerically greater response rates on the modified ACR 20 (mACR20; the modified ACR20 omits ESR and CRP because they are infrequently measured in clinical practice) than infliximab-treated patients (etanercept + MTX: 43%; etanercept monotherapy: 41%; infliximab + MTX: 35%; infliximab monotherapy: 26%; P: NR). Two other observational studies, one
based on the British Society for Rheumatology Biologics register (n = 2,711), the other on a multicenter Swedish cohort study (n = 949) reported similar findings. The Swedish study attributed the lower response rates of infliximab to a lack of adherence to therapy over the 3 year follow-up period.

A well conducted retrospective cohort study did not meet our eligibility criteria, nevertheless we are presenting findings because this study was the only one that compared radiographic progression between etanercept and infliximab. This population-based study determined erosion progression and joint space narrowing on 372 Swiss patients who were monitored through the Swiss Clinical Quality Management System. Combination therapies of infliximab and DMARDs and etanercept and DMARDs did not present statistically significant differences in progression of erosion (Ratingen score; data NR; P = 0.07) after a mean follow-up of 1.7 years. The combination of infliximab and DMARDs however, led to statistically significantly lower joint space narrowing than etanercept and DMARDs (data NR; P < 0.001). This difference, however, was not obvious when the analysis was limited to MTX as the concomitant DMARD. The combination of infliximab and MTX was statistically significantly more efficacious on all outcome measures than etanercept monotherapy (data NR).

**Indirect Head-Head Comparisons**

In addition, we conducted adjusted indirect comparisons based on our meta-analyses of placebo-controlled trials to compare the treatment effects of individual TIMs. We included data from published studies or from the CDER website. If data was sufficient, we conducted meta-analyses and adjusted indirect comparisons using ACR50 responses as outcome measures. For all analyses we used only data derived from study arms at or near the recommended dosage.

We chose ACR50 because a 50 percent improvement is likely to translate to a clinically significant improvement in health-related quality of life. For example, a patient with 12 swollen and 8 tender joints at baseline would need to have fewer than six swollen and four tender joints at the trial endpoint. This would be accompanied by at least a 50 percent improvement in at least three of the following five measures: the patient’s assessment of pain, the patient’s assessment of global disease activity, the physician’s assessment of global disease activity, the HAQ-Disability Index, and either a C-reactive protein (CRP) or sedimentation rate (Westergren erythrocyte sedimentation rate [WESR]).

The underlying assumption for adjusted indirect comparisons to be valid is that the relative efficacy of an intervention is consistent across included studies. Included TIM-studies primarily differ in study duration, disease duration, and concomitant treatments. Differences in study durations did not appear to
be a factor altering the effect size. We included only studies of more than 3 months of study duration, however we did not limit by sample size. Most RCTs reported the onset of significant responses between 4 and 8 weeks. Treatment responses were sustained up to 2 years in open-label extension studies. Sensitivity analyses based on different study durations did not substantially change the point estimates of the treatment effect. Likewise, sensitivity analyses excluding studies without concomitant MTX treatment, or studies on patients with early RA, did not substantially change the point estimate. One exception was the sensitivity analysis of infliximab where removing a study on patients with early RA substantially changed the effect size. However, it was unclear if this effect was attributable to true heterogeneity or to a lesser influence of random error in this large trial. Results presented below exclude this study. Overall, diagnostic criteria and eligibility criteria appeared to be sufficiently similar to make adjusted indirect comparisons a reasonable approach. However, given the small number of studies and the subsequent lack of precision, results should still be interpreted cautiously.

Results of adjusted indirect comparisons are depicted in Table 5 and Figure 1; corresponding forest plots for meta-analyses are presented in Appendix E. Findings suggest that no substantial differences exist among the efficacy of adalimumab, etanercept, and infliximab. However, given the wide confidence intervals, clinically significant differences cannot be excluded with certainty. Confidence intervals encompass differences that would be clinically significant. More data is needed to increase the precision of these estimates.

Point estimates favor adalimumab, etanercept, and infliximab over anakinra. However, differences do not reach statistical significance in adjusted indirect comparisons which is likely attributable to a lack of power. Adjusted indirect comparisons of anti-TNF drugs as a class compared to anakinra result in a statistically significantly greater efficacy of anti-TNF drugs on ACR 20 but not on ACR 50. Figure 1 depicts results of adjusted indirect comparisons of anakinra with adalimumab, etanercept, infliximab, and anti-TNF drugs as a class.

<table>
<thead>
<tr>
<th>Table 5. Adjusted Indirect Comparisons of TIMs for the Treatment of RA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comparison</strong></td>
</tr>
<tr>
<td>Adalimumab vs. Etanercept</td>
</tr>
<tr>
<td>Adalimumab vs. Infliximab</td>
</tr>
<tr>
<td>Anakinra vs. Adalimumab</td>
</tr>
<tr>
<td>Anakinra vs. Etanercept</td>
</tr>
<tr>
<td>Anakinra vs. Infliximab</td>
</tr>
<tr>
<td>Etanercept vs. Infliximab</td>
</tr>
</tbody>
</table>
Figure 1. Adjusted Indirect Comparisons of Anakinra with anti-TNF Drugs for the Treatment of RA

H. General Efficacy
Because of the lack of head-to-head trials, we reviewed placebo-controlled trials. We have summarized evidence on the general efficacy of TIMs in the treatment of RA. This, however, does not provide evidence on the comparative efficacy and tolerability of TIMs. If we identified high quality meta-analyses, we report the pooled estimates but do not describe the results of individual component studies, except when outcome measures of interest are reported (e.g., quality of life, functional capacity) that were not quantitatively analyzed in a meta-analysis.
**Abatacept**
Four trials examined the efficacy of abatacept in patients with RA (six publications).\(^{37-42}\) The largest study was a good multi-national trial enrolling 652 patients with MTX-resistant RA.\(^{38}\) After 1 year of follow-up, abatacept (10mg/kg) led to statistically significant improvements on all outcome measures (ACR20/50/70, HAQ-DI, DAS28, SF-36, Genant modified Sharp scores). At 1 year, 48.3 percent of abatacept- and 18.2 percent of placebo-treated patients achieved an ACR 50 response (\(P < 0.001\)), 28.8 percent versus 6.1 percent achieved an ACR 70 response (\(P < 0.001\)). Two phase II studies reported consistent findings.\(^{39-42}\)

A good 6-month trial was conducted in patients with an inadequate response to anti-TNF treatment (etanercept or infliximab).\(^{37}\) After 6 months of treatment, abatacept led to statistically significant improvement on all outcome measures compared to placebo (ACR20/50/70, DAS28, HAQ, SF-36).

**Adalimumab**
Six fair-rated studies examined the efficacy of adalimumab in patients with RA.\(^{31, 43-47}\) Five studies were conducted in patients who have failed standard DMARD therapies.\(^{43-47}\) Overall, 2,354 patients with active RA, not adequately responding to standard DMARD therapies, were included. In one study, participants remained on their standard antirheumatic therapy regardless of the DMARD therapy.\(^{43}\) Two trials allowed only MTX as a concomitant DMARD,\(^{44, 47}\) and in two studies no DMARDS were permitted as concomitant treatments.\(^{45, 46}\) The longest study lasted 52 weeks,\(^{44}\) study durations of the other trials were 12 weeks,\(^{45}\) 24 weeks,\(^{43, 47}\) and 26 weeks,\(^{46}\) respectively. The most common dosing regimen was 40 mg adalimumab biweekly; however, doses ranged from 20 mg and 40 mg weekly to 80 mg biweekly. Across all dosing regimens, response rates compared to placebo on ACR20/50/70 were significantly greater for adalimumab. Likewise, significantly more patients on adalimumab achieved improvements in health outcome measures (HAQ, SF-36, FACIT [Functional Assessment of Chronic Illness Therapy]) than patients on placebo. In the 52-week trial, 41.5 percent of patients on adalimumab 40 mg biweekly achieved an ACR50 response, compared to 9.5 percent on placebo (\(P < 0.001\)).\(^{44}\) HAQ scores at 52 weeks also significantly favored the adalimumab 40 mg biweekly group (-.59 vs. -.25; \(P < 0.001\)). The radiographic progression of disease as assessed on the modified Sharp score was significantly less in adalimumab-treated patients at study endpoint (\(P < 0.001\)).

The PREMIER study was the only trial conducted in MTX naïve patients with early, aggressive RA.\(^{31}\) This multinational study randomized 799 patients with early RA to a combination of adalimumab (40 mg every other week) and MTX (20mg / week), adalimumab monotherapy (40 mg every other week), or MTX monotherapy (20mg / week). After 1 year of follow-up, patients on the combination therapy...
exhibited statistically significantly greater responses on ACR 50 than patients on adalimumab and MTX monotherapies (62% vs. 41% vs. 46%; \( P < 0.001 \)). In addition, they had statistically significantly less progression on the modified Sharp score (1.9 vs 5.5 vs. 10.4 Sharp units; \( P < 0.002 \)). Forty-nine percent of patients on the combination therapy achieved remission (DAS28 < 2.6) after 2 years of treatment, compared with 23 percent on adalimumab monotherapy and 21 percent on MTX monotherapy (\( P < 0.001 \)).

We pooled data of the five studies described above to receive summary effect sizes for a treatment regimen of 40mg adalimumab biweekly, which is the recommended dosage for the treatment of RA. Our outcomes of choice were pooled relative risk (benefit) ratios to achieve ACR 20/50/70 responses and the corresponding NNTs. The NNTs (benefit) for ACR20/50/70 are 3 (95% CI 2-4), 4 (95% CI 3-6), and 8 (95% CI 6-11), respectively. In other words, three patients have to be treated with adalimumab to achieve one more ACR20 response than placebo; four patients to achieve an additional ACR50 response and eight patients for an additional ACR70 response. Because of moderate heterogeneity (\( I^2 \)-statistics), we used random effects models. The small number of component studies did not enable us to reliably assess publication bias. Reported data was not sufficient to calculate pooled estimates for HAQ. Study characteristics, pooled relative risk ratios, and forest plots are presented in Appendix E.

**Anakinra**

We identified one high quality meta-analysis that pooled one unpublished and three published RCTs. Overall, this Health Technology Assessment from the United Kingdom (UK) included 1,007 patients. Pooled results presented statistically significantly greater improvements of anakinra- than placebo-treated patients on all outcome measures (ACR20/50/70, HAQ, Patient Global Assessment). The NNTs to achieve one additional responder on ACR20/50/70 were 7, 11, and 33, respectively. Adjusted indirect comparisons with two anti-TNF agents (etanercept, infliximab) suggested that anakinra may be significantly less effective at relieving clinical symptoms than anti-TNF drugs (ACR20: RR 0.21; 95%CI 0.10-0.32). We replicated this indirect comparison with a larger number of studies assessing anti-TNF drugs. Although our results also suggest that anakinra is significantly less effective in achieving an ACR20 response than TNF inhibitors as a class, the effect size was smaller in our calculations than in the results of the UK report and just reached statistical significance (RR: 0.67; 95%CI 0.45-0.99). Furthermore, indirect comparisons of ACR50 response rates did not present a statistically significant difference (RR: 0.69; 95%CI 0.39-1.22) Corresponding forest plots are presented in Appendix E.
A fair RCT, not included in the meta-analysis described above, reported similar results for patients with active RA who were treated with 100 mg anakinra or placebo for 24 weeks. Anakinra had significantly higher response rates than placebo (ACR50: 17% vs. 8%; \( P < 0.01 \)) and faired significantly better on all health outcome measures (HAQ: -0.29 vs. -0.18; \( P < 0.05 \); patient’s assessment of disease activity: -17.7 vs. -8.9; \( P < 0.001 \); patient’s assessment of pain: -19.0 vs. -11.7; \( P < 0.01 \)).

We pooled data from three trials that provided sufficient information for critical, methodological appraisal. We did not include a study that was published as an abstract only. Our outcomes of choice were pooled relative risk (benefit) ratios to achieve ACR 20/50/70 responses and the corresponding NNTs. Because of moderate heterogeneity (I\(^2\)-statistics), we used random effects models. The NNTs (benefit) for ACR20/50/70 are 6 (95%CI 4-9), 10 (95%CI 7-18), and 35 (95%CI 75[harm]-14[benefit]) respectively. In other words, six patients have to be treated with anakinra to achieve one more ACR20 response than placebo; 10 patients to achieve an additional ACR50 response and 35 patients for an additional ACR70 response. The NNT for an ACR70 response did not reach statistical significance and thus the confidence interval includes the possibility of harm. The small number of component studies did not enable us to reliably assess publication bias. Reported data was not sufficient to calculate pooled estimates for HAQ. Study characteristics, pooled relative risk ratios, and forest plots are presented in Appendix E.

**Etanercept**

Two well conducted meta-analyses examined the efficacy of etanercept in patients with RA. Both studies reported significantly greater improvements for etanercept-treated patients at study endpoint. Pooled results indicated that 39 percent of patients treated with the recommended dose of 50 mg etanercept per week reached an ACR50 response, compared to four percent of patients on placebo (RR: 8.89; 95% CI 3.61 – 21.89). The NNT to achieve one additional ACR50 response was 3.

Two fair trials compared etanercept to MTX over 52 weeks. Although both studies failed to show statistically significant differences between etanercept (25 mg twice weekly) and MTX (20 mg/week) in health outcome measures (SF-36, HAQ, ASHI), and ACR response rates at study endpoints (52 weeks), radiographic outcomes were significantly better in patients on ETA than on MTX. Improved radiographic outcomes were maintained during an extension of the ERA (Early Rheumatoid Arthritis) trial to 24 months. Both trials report statistically significantly better efficacy outcomes for etanercept- than for MTX-treated patients during the first months of treatment. One study was conducted in patients with early RA. The TEMPO (Trial of Etanercept and Methotrexate with Radiographic Patient Outcomes) study,
which was conducted in 686 patients with moderate to severe RA, provided similar results on health outcomes.\textsuperscript{29, 51} In addition, this study compared etanercept and MTX mono-therapies to a combination of MTX (20 mg/week) and etanercept (25 mg twice weekly). Overall, the combination treatment achieved significantly better results on most outcome measures than etanercept and MTX alone. A significantly higher proportion of patients on the combination treatment than on MTX and etanercept reached ACR50 response after 52 weeks (69% vs. 43%; 69% vs. 48%; \(P < 0.0001\) for both comparisons) or were on remission (DAS < 1.6; 35% vs. 13%; 35% vs. 16%; \(P < 0.0001\) for both comparisons). Likewise, HAQ scores were significantly higher for the combination treatment than for either of the monotherapies (\(P < 0.001\)).\textsuperscript{51} Patients on the combination treatment presented a significantly greater retardation of joint damage than patients on MTX or etanercept monotherapy. This study reported no differences in adverse events.

A fair, 12-week trial assessed health-related quality of life as a secondary outcome measure (HAQ, SF-36, feeling thermometer) in patients with longstanding RA who had failed DMARD treatments.\textsuperscript{53, 54} Two regimens of etanercept (10 mg and 25 mg twice weekly) were compared to placebo; no DMARDs were allowed. Both etanercept groups achieved statistically significantly greater improvements on all outcome measures compared to placebo.

A fair, 24-week study did not detect any synergistic effects of a combination treatment of etanercept (25 mg or 50 mg/week) and anakinra (100 mg/day) compared to etanercept monotherapy.\textsuperscript{32} Overall, 242 patients who were on stable doses of MTX treatment were enrolled. At endpoint, combination treatment did not lead to greater efficacy than etanercept only. Furthermore, the frequency of serious adverse events was substantially higher in the combination groups (14.8% for 50 mg etanercept + anakinra, 4.9% for 25 mg etanercept + anakinra vs. 2.5% for etanercept only; no P-values reported). Likewise, withdrawals due to adverse events were higher in the combination groups than in the etanercept group (8.6% vs. 7.4%; no P-values reported).

We pooled data from five studies\textsuperscript{29, 54, 65-67} to receive summary effect sizes for a treatment regimen of 50 mg etanercept per week, which is the recommended dosage for the treatment of RA. Our outcomes of choice were pooled relative risk (benefit) ratios to achieve ACR20/50/70 responses and the corresponding NNTs. Because of high heterogeneity (\(I^2\)-statistics), we used random effects models. The high heterogeneity was mainly attributable to the Klareskog et al.\textsuperscript{29} study, which was larger and of higher methodological quality than the remaining studies. Effect sizes in this study were smaller than in the other studies. No substantial differences in study populations, concomitant treatments, or study durations could
explain the high heterogeneity. The most likely explanation is the small number of component studies and
the higher methodological quality of the Klareskog et al. study. The directionality of the treatment effect
is consistent for all studies and favors etanercept. The NNTs (benefit) for ACR20/50/70 were 2 (95%CI
1-5), 3 (95%CI 2-4), and 5 (95%CI 4-8), respectively. In other words, two patients have to be treated with
etanercept to achieve one more ACR20 response than placebo; three patients to achieve an additional
ACR50 response and eight patients for an additional ACR70 response. The small number of component
studies did not enable us to reliably assess publication bias. Reported data was not sufficient to calculate
pooled estimates for HAQ. Study characteristics, pooled relative risk ratios, and forest plots are presented
in Appendix E.

Infliximab
Two well conducted meta-analyses determined the general efficacy of infliximab in RA. Pooled
results of both studies report significantly greater improvements on all outcome measures than placebo.
For 10 mg infliximab every 8 weeks, the ACR50 response rate was 30 percent compared to 5 percent for
placebo. The NNT to achieve one additional response was 4. Two recent studies were not included in
these meta-analyses.

The ASPIRE (Active-controlled Study of Patients Receiving Infliximab for the Treatment of Rheumatoid
Arthritis of Early Onset) enrolled 1,049 patients with early RA and compared the benefits of initiating
treatment with MTX (20 mg) alone or a combination of MTX and infliximab (3 mg/kg or 6 mg/kg) over
52 weeks. At endpoint, patients in the combination groups had significantly higher ACR-N (ACR
composite score) improvements than patients on MTX monotherapy (38.9% [3 mg infliximab] vs. 46.7%
[6 mg infliximab] vs. 26.4% [placebo]; P < 0.001); the ACR50 response was 45.6% vs. 40.4% vs. 32.1%,
respectively. In addition, HAQ and SF-36 scores improved significantly more in the combination groups
than in the MTX group. More patients in the combination groups had serious adverse events (14% vs.
11%; P-value not reported) and serious infections (5.6% [3 mg/kg infliximab] vs. 5.0% [6 mg/kg
infliximab] vs. 2.1% [MTX]; P = 0.02 and P = 0.04) than patients on placebo. Results of an open-label
extension of a 52-week RCT included in one of the meta-analyses reported that response rates on HAQ
and SF-36 were maintained for another year. Radiographic progression of disease was significantly
lower than in the MTX only group. Patients on the combination treatment also had a higher probability of
maintaining their employability compared with those on MTX alone. A smaller (n = 147), fair RCT
conducted in Japanese patients provided similar results.
The START (Trial for Rheumatoid Arthritis with Remicade) study was designed to assess the risk of serious infections, however, it also determined the response rates of infliximab treatment in combination with ongoing rheumatic therapy compared to placebo.\textsuperscript{57} This multinational study enrolled 1084 patients with RA applying minimal exclusion criteria. Patients received a fixed dose regimen of 3mg/kg or 10mg/kg infliximab for 22 weeks. At endpoint, patients on both regimens achieved statistically significantly greater response rates on all outcome measures (ACR20/50/70, DAS-28) than patients on placebo.

We pooled data from six studies\textsuperscript{55-57, 68, 71} to receive summary effect sizes for a treatment regimen of 3-10 mg/kg infliximab every 4 to 8 weeks, which is the recommended dosage for the treatment of RA. Our outcomes of choice were pooled relative risk (benefit) ratios to achieve ACR 20/50/70 responses and the corresponding NNTs. We assumed that Paulus response rates are very similar to ACR response rates. Because of high heterogeneity ($F$-statistics), we used random effects models. The high heterogeneity was mainly attributable to the St. Clair et al. study,\textsuperscript{56} which was larger and conducted in MTX naïve patients with early RA. Effect sizes in this study were smaller than in the other studies. In a sensitivity analysis we removed the St. Clair et al. study, which substantially reduced heterogeneity. Because it is unclear if the smaller treatment effect in the St.Clair et al. study is attributable to less random error in this large study or to true heterogeneity, we present the pooled relative risks with and without St. Clair et al. in Appendix E. Data was not sufficient to pool for ACR70 response rates. The small number of component studies did not enable us to reliably assess publication bias. Reported data was not sufficient to calculate pooled estimates for HAQ. The NNTs (benefit) for ACR20/50 (without St. Clair et al.) was 2.9 (95%CI 2.6 -3.3) and 4.2 (95%CI 3.7-4.8). In other words, three patients have to be treated with infliximab to achieve one more ACR20 response than placebo; four patients to achieve an additional ACR50 response. NNTs were identical for estimates including the St. Clair et al. study. Study characteristics, pooled relative risk ratios, and forest plots are presented in Appendix E.

**Rituximab**

Two fair, 24-week studies assessed the general efficacy of rituximab for the treatment of patients with DMARD resistant RA.\textsuperscript{58, 59} Both trials reported statistically significantly greater response rates for rituximab treated- than placebo treated patients. In the larger trial (n = 465), rituximab regimens (2 x 500mg or 2 1000mg) led to statistically significantly greater response rates on ACR 20 than placebo (55% vs. 54% vs. 28%; $P < 0.0001$).\textsuperscript{59} Likewise, patients on rituximab achieved statistically significantly greater responses on ACR50 (data NR; $P < 0.001$) and ACR 70 (data NR; $P < 0.001$) The primary efficacy analysis, however, was limited to RF – positive (rheumatoid factor – positive) patients. Adding
RF-negative patients to the efficacy analysis limited the statistically significant differences to the 1000mg regimen (data NR). Furthermore, postrandomization exclusions reduce the internal validity of this study. Both studies did not assess radiographic outcomes.
### Table 6. Summary of Efficacy Trials in Adult Patients with RA

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Comparisons</th>
<th>Primary outcome</th>
<th>Secondary outcomes</th>
<th>Population</th>
<th>Results</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geborek et al. 2002(^{24})</td>
<td>Non-randomized trial</td>
<td>404</td>
<td>12 months</td>
<td>ETA/ INF/ Leflunomide</td>
<td>ACR20/50</td>
<td>DAS28</td>
<td>Population-based; active RA; had failed at least one DMARD treatment; mean disease duration: 14.5 yrs.</td>
<td>ACR 20 response rates significantly greater for ETA than for INF at 3 months ($P&lt;0.02$) and 6 months ($P&lt;0.05$); no differences at 12 months</td>
<td>Fair</td>
</tr>
<tr>
<td>Finckh et al. 2006(^{33})</td>
<td>Retrospective cohort study</td>
<td>372</td>
<td>1.7 years</td>
<td>ETA/ ETA + MTX/ INF+MTX</td>
<td>Erosion progression</td>
<td>Joint space narrowing</td>
<td>Population-based; active RA; patients who have started a biologic</td>
<td>No difference in erosion progression between ETA + MTX and INF + MTX. Significantly better outcomes for INF + MTX than for ETA monotherapy</td>
<td>N/A</td>
</tr>
<tr>
<td>Hyrich et al, 2006(^{61})</td>
<td>Prospective cohort study</td>
<td>2,711</td>
<td>6 months</td>
<td>ETA, INF</td>
<td>EULAR</td>
<td>DAS 28</td>
<td>Population-based; active RA; started a biologic; mean disease duration: 14.6 yrs.</td>
<td>EULAR response rates numerically greater for ETA than for INF at 6 months</td>
<td>Fair</td>
</tr>
<tr>
<td>Kristensen et al. 2006(^{26})</td>
<td>Prospective cohort study</td>
<td>949</td>
<td>3 years</td>
<td>ETA, INF</td>
<td>EULAR</td>
<td>ACR20/50/70</td>
<td>Population-based; active RA; started a biologic; mean disease duration: 13.4 yrs.</td>
<td>Moderate EULAR and ACR response rates numerically greater for ETA than for INF at 3 years</td>
<td>Fair</td>
</tr>
<tr>
<td>Weaver et al. 2006(^{25})</td>
<td>Prospective cohort study</td>
<td>3694</td>
<td>12 months</td>
<td>ETA, INF</td>
<td>mACR20</td>
<td>HAQ</td>
<td>Primary-care based; active RA; patients who needed change in treatment regimen; mean disease duration: NR</td>
<td>mACR 20 response rates numerically greater for ETA than for INF at 12 months</td>
<td>Fair</td>
</tr>
</tbody>
</table>

### ABATACEPT

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Comparisons</th>
<th>Primary outcome</th>
<th>Secondary outcomes</th>
<th>Population</th>
<th>Results</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genovese et al. 2005(^{37})</td>
<td>RCT</td>
<td>393</td>
<td>6 months</td>
<td>ABA + DMARD or AKA / placebo + DMARD or AKA</td>
<td>ACR 20, HAQ</td>
<td>DAS28, ACR50/70 SF-36</td>
<td>Patients who had an inadequate response to etanercept or infliximab; mean disease duration: 11.9 yrs.</td>
<td>Statistically significantly greater improvements on all outcome measures for ABA</td>
<td>Good</td>
</tr>
<tr>
<td>Study Authors and Year</td>
<td>Design</td>
<td>n</td>
<td>Study Duration</td>
<td>Treatment</td>
<td>End Points</td>
<td>Exposure</td>
<td>Outcome</td>
<td>Comments</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>--------</td>
<td>----</td>
<td>----------------</td>
<td>-----------</td>
<td>------------</td>
<td>----------</td>
<td>---------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>Kremer et al., 2006(^{38})</td>
<td>RCT</td>
<td>652</td>
<td>12 months</td>
<td>ABA + MTX / Placebo + MTX</td>
<td>ACR 20</td>
<td>HAQ-DI, ACR50/70, radiographic evaluation</td>
<td>Active RA for at least 1 year; had failed MTX treatment; mean disease duration: 8.7 yrs.</td>
<td>Statistically significantly greater improvements on all outcome measures for ABA</td>
<td>Fair</td>
</tr>
<tr>
<td>Kremer et al., 2005(^{39,40,41})</td>
<td>RCT</td>
<td>339</td>
<td>12 months</td>
<td>ABA + MTX / Placebo + MTX</td>
<td>ACR 20</td>
<td>ACR50/70 DAS28, HAQ</td>
<td>Active RA for at least 6 months with a stable dose of MTX; mean disease duration: 9.4 yrs.</td>
<td>Statistically significantly greater improvements on all outcome measures for ABA</td>
<td>Fair</td>
</tr>
<tr>
<td>Moreland et al. 2002(^{42})</td>
<td>RCT</td>
<td>122</td>
<td>12 weeks</td>
<td>ABA + DMARD / placebo + DMARD</td>
<td>ACR 20</td>
<td>ACR50/70</td>
<td>Active RA for less than 7 years; failed at least 1 DMARD treatment; mean disease duration: 3.4 yrs.</td>
<td>Numerically greater improvements on ACR core data for ABA</td>
<td>Fair</td>
</tr>
<tr>
<td><strong>ADALIMUMAB</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breedveld et al. 2006 (PREMIER)(^{31})</td>
<td>RCT</td>
<td>799</td>
<td>2 years</td>
<td>ADA + placebo / ADA + MTX / MTX + placebo</td>
<td>ACR 50, modified Sharp score</td>
<td>ACR20/70/90, DAS28, HAQ,</td>
<td>Early, aggressive RA; MTX naïve patients; mean disease duration: 0.7 yrs.</td>
<td>Combination of ADA + MTX had statistically significantly greater improvements on ACR 50</td>
<td>Good</td>
</tr>
<tr>
<td>Furst et al. 2003(^{43})</td>
<td>RCT</td>
<td>636</td>
<td>24 weeks</td>
<td>ADA +Standard RA therapy/ Placebo + Standard RA therapy</td>
<td>safety</td>
<td>ACR20/50/70, HAQ</td>
<td>Active RA for at least 3 months; DMARD naïve/or on stable regimen; mean dis. duration: 10.5 yrs.</td>
<td>ACR20/50/70 response rates significantly greater with ADA than with placebo</td>
<td>Fair</td>
</tr>
<tr>
<td>Keystone et al. 2004(^{44})</td>
<td>RCT</td>
<td>619</td>
<td>52 weeks</td>
<td>ADA +MTX/ Placebo + MTX</td>
<td>Sharp, ACR 20, HAQ</td>
<td>ACR 50/70</td>
<td>Active RA; on stable MTX regimen; mean disease duration: 11 yrs.</td>
<td>ACR20/50/70 response rates significantly greater with ADA than with placebo</td>
<td>Fair</td>
</tr>
<tr>
<td>Van de Putte et al. 2003(^{45})</td>
<td>RCT</td>
<td>284</td>
<td>12 weeks</td>
<td>ADA/ Placebo</td>
<td>ACR 20</td>
<td>ACR50; ACR70; TJC; SJC; DAS28; HAQ.</td>
<td>Active RA; had failed at least one DMARD treatment; mean disease duration: 10 yrs.</td>
<td>ACR20/50/70 response rates significantly greater with ADA than with placebo</td>
<td>Fair</td>
</tr>
<tr>
<td>Van de Putte et al. 2004(^{46})</td>
<td>RCT</td>
<td>544</td>
<td>26 weeks</td>
<td>ADA / Placebo</td>
<td>ACR20</td>
<td>ACR50/70, HAQ</td>
<td>Active RA; had failed at least one DMARD treatment; mean dis.</td>
<td>ACR20/50/70 response rates significantly greater with ADA than with placebo</td>
<td>Fair</td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>n</td>
<td>Duration</td>
<td>Intervention</td>
<td>Outcomes</td>
<td>Comparator</td>
<td>Conclusion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>----</td>
<td>----------</td>
<td>--------------</td>
<td>----------</td>
<td>------------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weinblatt et al. 2003&lt;sup&gt;47&lt;/sup&gt;</td>
<td>RCT</td>
<td>271</td>
<td>24 weeks</td>
<td>ADA+MTX / MTX + Placebo</td>
<td>ACR20, HAQ</td>
<td>ACR 50/70, SF-36</td>
<td>Active RA; on stable MTX regimen; had failed at least one other DMARD; mean disease duration: 11 yrs.</td>
<td>ACR20/50/70 response rates significantly greater with ADA than with placebo</td>
<td>Fair</td>
</tr>
<tr>
<td>Clark et al. 2004&lt;sup&gt;27&lt;/sup&gt;</td>
<td>MA</td>
<td>1007</td>
<td>&gt; 6 mo</td>
<td>MTX + Placebo</td>
<td>ACR20/50/70</td>
<td>HAQ</td>
<td>Adults with RA</td>
<td>ACR20/50/70 response rates significantly greater with AKA than with placebo; adjusted indirect comparisons suggest that AKA is significantly less efficacious than anti-TNF</td>
<td>Good</td>
</tr>
<tr>
<td>Cohen et al. 2004&lt;sup&gt;48&lt;/sup&gt;</td>
<td>RCT</td>
<td>501</td>
<td>24 weeks</td>
<td>AKA+MTX/ MTX+Placebo</td>
<td>ACR20</td>
<td>ACR50/70, HAQ</td>
<td>&gt; 6 months history of active RA; stable MTX regimen; mean disease duration: 10.5 yrs.</td>
<td>ACR20/50/70 response rates at 24 weeks significantly greater with AKA than with placebo</td>
<td>Fair</td>
</tr>
<tr>
<td>Blumenauer et al. 2003&lt;sup&gt;34&lt;/sup&gt;</td>
<td>MA</td>
<td>955</td>
<td>&gt; 6 mo</td>
<td>ETA(+MTX)/ (MTX+) placebo</td>
<td>ACR20/50/70</td>
<td>safety</td>
<td>Adults with RA</td>
<td>ACR20/50/70 response rates significantly greater with ETA than with placebo</td>
<td>Good</td>
</tr>
<tr>
<td>Jobanputra et al. 2002&lt;sup&gt;15&lt;/sup&gt;</td>
<td>MA</td>
<td>1062</td>
<td>4 weeks – 1 year</td>
<td>ETA(+MTX) / (MTX +)placebo</td>
<td>ACR20/50/70</td>
<td>safety</td>
<td>Adults with RA</td>
<td>ACR20/50/70 response rates significantly greater with ETA than with placebo</td>
<td>Good</td>
</tr>
<tr>
<td>Bathon et al. 2000&lt;sup&gt;28, 30, 49&lt;/sup&gt;</td>
<td>RCT</td>
<td>632</td>
<td>52 weeks</td>
<td>ETA / MTX</td>
<td>ACR20/50/70</td>
<td>SF-36, HAQ, ACR-N, modified Sharp</td>
<td>early, active RA; mean disease duration: 1 yr.</td>
<td>Up to 6 months significantly higher ACR 50/70 response rates for ETA than for MTX; no differences after. At 12 months no differences in ACR20 but less joint erosion for ETA; no significant differences in SF-36, HAQ, and ASHI</td>
<td>Fair</td>
</tr>
<tr>
<td>study</td>
<td>study type</td>
<td>n</td>
<td>duration</td>
<td>treatment</td>
<td>outcome measures</td>
<td>criteria</td>
<td>quality assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>------------</td>
<td>-----</td>
<td>-----------</td>
<td>-----------</td>
<td>---------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
<td>-------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genovese et al. 2004 (ETA + MTX)</td>
<td>RCT</td>
<td>242</td>
<td>24 weeks</td>
<td>ETA + MTX / ETA + AKA + MTX</td>
<td>ACR50, ACR20/70, SF-36</td>
<td>&gt; 6 months history of active RA; stable MTX regimen; mean disease duration: 10 yrs.</td>
<td>No additional benefit from ETA-AKA combination therapy; Adverse event rates significantly higher in combination than in ETA group</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Klareskog et al. 2004 (TEMPO)</td>
<td>RCT</td>
<td>682</td>
<td>52 weeks</td>
<td>ETA / MTX / MTX + ETA</td>
<td>Sharp, ACR20/50/70, HAQ</td>
<td>&gt; 6 months history of active RA; unsatisfactory response to at least one DMARD other than MTX; mean disease duration: 6.6 yrs.</td>
<td>ETA + MTX regimen achieved better results on most outcome measures than ETA or MTX monotherapies</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Moreland et al. 1999 (MTX)</td>
<td>RCT</td>
<td>234</td>
<td>26 weeks</td>
<td>ETA / Placebo</td>
<td>ACR20/50, SF-36, HAQ</td>
<td>Active RA; had failed 1 to 4 DMARD treatments other than MTX; mean disease duration: 12 yrs.</td>
<td>ACR20/50 response rates, HAQ and SF-36 scores significantly greater with ETA than with placebo</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Abe et al. 2006 (INF)</td>
<td>RCT</td>
<td>147</td>
<td>14 weeks</td>
<td>MTX + placebo</td>
<td>ACR20/50/70, Safety</td>
<td>&gt; 6 months history of active RA; stable MTX regimen; mean disease duration: 7.9 yrs.</td>
<td>ACR20/50/70 response rates at 14 weeks significantly greater with INF than with placebo</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Blumenauer et al. 2002 (INF)</td>
<td>MA</td>
<td>529</td>
<td>&gt; 6mo</td>
<td>MTX + Placebo</td>
<td>ACR20/50/70, Withdrawals, safety</td>
<td>Adults with RA</td>
<td>ACR20/50/70 response rates significantly greater with INF than with placebo</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Jobanputra et al. 2002 (INF)</td>
<td>MA</td>
<td>630</td>
<td>4 weeks – 1 year</td>
<td>MTX + Placebo</td>
<td>ACR20/50/70, Safety</td>
<td>Adults with RA</td>
<td>ACR20/50/70 response rates significantly greater with INF than with placebo</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>St. Clair et al. 2004 (INF)</td>
<td>RCT</td>
<td>1049</td>
<td>54 weeks</td>
<td>INF + MTX / MTX</td>
<td>ACR-N, ACR20/50/70, Sharp</td>
<td>Early RA, MTX naïve patients; mean disease duration: 0.9</td>
<td>ACR20/50/70 response rates and HAQ scores were significantly greater with INF than with placebo</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Design</td>
<td>N</td>
<td>Duration</td>
<td>Treatment</td>
<td>End Point</td>
<td>Comparator</td>
<td>Outcome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>----</td>
<td>----------</td>
<td>-----------</td>
<td>-----------</td>
<td>------------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Westhovens et al., 2006 (START)&lt;sup&gt;57&lt;/sup&gt;</td>
<td>RCT</td>
<td>1084</td>
<td>22 weeks</td>
<td>INF + MTX / MTX</td>
<td>Safety</td>
<td>ACR20/50/70, DAS-28</td>
<td>INF+MTX than with MTX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edwards et al. 2004&lt;sup&gt;58&lt;/sup&gt;</td>
<td>RCT</td>
<td>161</td>
<td>24 weeks</td>
<td>RIT + MTX / RIT + placebo / RIT + cyclophosphamide / MTX + placebo</td>
<td>ACR50</td>
<td>ACR20/70, DAS28</td>
<td>INF+MTX than with MTX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emery et al. 2006&lt;sup&gt;59&lt;/sup&gt;</td>
<td>RCT</td>
<td>465</td>
<td>24 weeks</td>
<td>RIT (500mg) + MTX / RIT (1000mg) + MTX / MTX + placebo</td>
<td>ACR50</td>
<td>ACR20/70, DAS28</td>
<td>INF+MTX than with MTX</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**ACR20/50/70 response rates and DAS-28 scores were significantly greater with INF+MTX than with MTX.**

**RITUXIMAB**

- **Westhovens et al., 2006 (START)<sup>57</sup>**
  - N: 1084
  - Duration: 22 weeks
  - Treatment: INF + MTX / MTX
  - End Point: Safety
  - Comparator: ACR20/50/70, DAS-28
  - Outcome: INF+MTX than with MTX

- **Edwards et al. 2004<sup>58</sup>**
  - N: 161
  - Duration: 24 weeks
  - Treatment: RIT + MTX / RIT + placebo / RIT + cyclophosphamide / MTX + placebo
  - End Point: ACR50
  - Comparator: ACR20/70, DAS28
  - Outcome: INF+MTX than with MTX

- **Emery et al. 2006<sup>59</sup>**
  - N: 465
  - Duration: 24 weeks
  - Treatment: RIT (500mg) + MTX / RIT (1000mg) + MTX / MTX + placebo
  - End Point: ACR50
  - Comparator: ACR20/70, DAS28
  - Outcome: INF+MTX than with MTX

**Targeted Immune Modulators**

ABA: abatacept; ACR 20/50/70: American College of Rheumatology, numbers refer to percentage improvement; ADA: adalimumab; AKA: anakinra; ASHI: arthritis-specific health index; DAS28: disease activity score; DMARD: disease-modifying antirheumatic drug; ETA: etanercept; EULAR: European League Against Rheumatism; HAQ: Health Assessment Questionnaire; INF: infliximab; MA: meta-analysis; MTX: methotrexate; RA: rheumatoid arthritis; RCT: randomized controlled trial; RIT: rituximab; SF-36: Medical Outcomes Study Short Form 36 Health Survey; SJC: swollen joint count; TJC: tender joint count.
II. Juvenile Rheumatoid Arthritis (JRA)
Currently only etanercept is approved by the FDA for the treatment of JRA.

A. Summary of the Evidence
The evidence on the comparative effectiveness of TIMs for the treatment of JRA is poor. One RCT provides fair evidence that etanercept is more efficacious than placebo for the treatment of JRA. However, the highly selected study population is likely to compromise the external validity of this study. One uncontrolled study does not provide convincing evidence on the generally efficacy of infliximab.

B. Description of Studies
For JRA, we did not find any head-to-head trials that compared one TIM to another. We found one placebo-controlled RCT with a 3-month, uncontrolled, open-label run-in phase assessing the efficacy and safety of etanercept. In addition, we included a retrospective analysis of data from a German registry for treatment of JRA and one small, uncontrolled, open-label trial on infliximab that we determined to be of poor quality. We did not detect any studies on abatacept, adalimumab, alefacept, anakinra, efalizumab, and rituximab. Included studies are presented in Table 7.

C. Study Population
Patients in the trials suffered from active polyarticular JRA and were between 4 and 17 years of age. Patients had active disease despite treatment with corticosteroids and MTX. Patients with concurrent medical conditions were excluded. The observational study included data of children with juvenile idiopathic arthritis, regardless of the subtype.

D. Outcome Measures
Response based on the Giannini criteria was the primary outcome measure for the open-label trial and the retrospective analysis. The primary outcome measure in the RCT was the number of patients with disease flare. It is unclear if this assessment was based on a validated rating scale. Additional outcome measures were the articular severity score, duration of morning stiffness, degree of pain, and CRP. The uncontrolled infliximab trial also assessed functional disability (HAQ) and health-related quality of life (SF-36).
E. Methodological Quality
In the etanercept study, only patients who had responded to etanercept treatment during a 3-month open-label run-in period were eligible for randomization (51 out of 69 patients). Therefore, the generalizability of findings will be low and results are likely to overestimate the true treatment effect and underestimate the incidence of adverse events. The infliximab study had fatal methodological flaws.

F. Sponsorship
Two studies were funded by the pharmaceutical industry. The RCT was also supported by the National Institute of Health. The funding of the infliximab study could not be determined.

G. Comparative Efficacy and Effectiveness
We did not identify any head-to-head trials.

H. General Efficacy
Because of the lack of head-to-head trials, we reviewed placebo-controlled trials. We have summarized evidence on the general efficacy of TIMs in the treatment of JRA.

Etanercept
Fifty-one patients were randomly assigned to etanercept (0.4 mg/kg twice weekly) or placebo. Study duration was 4 months. Significantly more patients on placebo than on etanercept had a disease flare (81% vs. 28%; \( P < 0.003 \)) during the study period. The median time to flare was 116 days for etanercept- and 28 days for placebo- treated patients (\( P < 0.001 \)). As stated above, the highly selected population is likely to have lead to an overestimation of the treatment effects. During the 3 month open-label run-in phase, 64 percent of patients achieved a 50 percent improvement of symptoms based on the Gianinni criteria. This response rate is comparable to that of a retrospective analysis of data of 322 patients treated with etanercept from a German registry. Sixty-one percent had a 50 percent improvement of symptoms at 3 months, 72 percent at 6 months. However, patients in this analysis were not limited to polyarticular JRA. The mean length of treatment in this study was 13.4 months. At one year, 82 percent of the non-systemic patients presented a 50 percent improvement. Subgroup analysis showed markedly lower response rates in patients with systemic arthritis.
**Infliximab**

One poor, uncontrolled study did not provide convincing evidence on the general efficacy of infliximab for the treatment of JRA. This uncontrolled open-label trial enrolled 24 females with polyarticular JRA. Sixty-two percent of patients dropped out during the first year, 17 percent because of infusion reactions. Completers-only analysis at one year reports significant improvements on clinical outcomes such as swollen or painful joints. However, neither HAQ nor SF-36 presented a statistically significant improvement at 1 year.
Table 7. Summary of Efficacy Trials in Patients with JRA

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Comparisons</th>
<th>Primary outcome</th>
<th>Secondary outcomes</th>
<th>Population</th>
<th>Results</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horneff et al. 2004⁷⁴</td>
<td>Retrospective data analysis</td>
<td>322</td>
<td>13.4 months</td>
<td>None</td>
<td>Response based on Gianinni criteria; Tolerability</td>
<td>Number of tender and swollen joints decreased during 3 months of treatment.</td>
<td>Active juvenile idiopathic arthritis; had failed at least one DMARD; mean disease duration: NR</td>
<td>Number of tender and swollen joints significantly decreased during 3 months of treatment.</td>
<td>N/A</td>
</tr>
<tr>
<td>Lovell et al. 2000³⁷³</td>
<td>Uncontrolled open-label trial / RCT</td>
<td>51</td>
<td>4 months</td>
<td>ETA / Placebo</td>
<td>Response based on Gianinni criteria; number of patients with disease flare Articular severity score, pain, CRP</td>
<td>Active polyartricular JRA; had failed corticosteroid and MTX treatment; mean disease duration: 5.8 yrs.</td>
<td>Significantly more patients on ETA than on placebo achieved 50% improvement</td>
<td>Fair</td>
<td></td>
</tr>
</tbody>
</table>

ADA: adalimumab; AKA: anakinra; DMARD: disease-modifying antirheumatic drug; ETA: etanercept; INF: infliximab; MA: meta-analysis; MTX: methotrexate
III. Ankylosing spondylitis (AS)
The following drugs are currently approved by the FDA for the treatment of AS: adalimumab, etanercept and infliximab.

A. Summary of the Evidence
Overall, the evidence on the comparative effectiveness of TIMs for the treatment of AS is poor. Good to fair evidence from five RCTs exists that etanercept and infliximab are significantly more efficacious than placebo for the treatment of AS. Treatment effects are large and consistent across studies. However, significant differences in study characteristics make this evidence insufficient to identify differences in efficacy among TIMs.

B. Description of Studies
For AS, we did not find any head-to-head trials comparing one TIM to another. We found five placebo-controlled trials; three trials assessed the efficacy of etanercept,76-78 two the efficacy of infliximab.79, 80 We did not detect any studies on abatacept, adalimumab, alefacept, anakinra, efalizumab, and rituximab. Included studies are presented in Table 8.

C. Study Populations
All patients suffered from active AS and were diagnosed based on the modified New York criteria.81 Disease duration and concomitant treatments varied across studies. Most patients used NSAIDS in addition to the study medication. The etanercept trials allowed corticosteroids and DMARDs as concomitant treatments.76-78 Patients in the infliximab trials were permitted to take only NSAIDS in addition to the study drug.79, 80 One study examined the efficacy of infliximab in patients with severe AS.79 Patients with an autoimmune disease other than AS, spinal fusion, a history of active listeriosis or mycobacterial infection, or recent antibiotic treatment were generally excluded from studies.

D. Outcome Measures
Most trials assessed response rates as defined by the Assessments in Ankylosing Spondylitis Working Group (ASAS).82 This scale (ASAS20/50/70 [figures refer to percentage improvement]), combines measures of global disease activity with functional capacity, pain, and acute phase laboratory parameters (see Appendix D). In addition, the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) was frequently assessed. Two studies evaluated health outcomes.77, 79
E. Methodological Quality
Study quality varied; one study was rated good,78 four were rated fair.76, 77, 79, 80 These “fair” ratings, however, are probably more attributable to inadequate reporting than to methodological flaws. Randomization methods and blinding were generally adequate; all studies used a double-dummy design (i.e., using an identical container for active treatment and placebo) to guarantee blinding. A high incidence of injection site reactions among users of etanercept de facto often overthrew blinding efforts.

F. Sponsorship
All trials were funded by the pharmaceutical industry.

G. Comparative Efficacy and Effectiveness
We did not identify any head-to-head trials.

H. General Efficacy
Because of the lack of head-to-head trials, we reviewed placebo-controlled trials. We have summarized evidence on the general efficacy of TIMs in the treatment of AS. This, however, does not provide evidence on the comparative efficacy and tolerability of TIMs.

Etanercept
One good78 and two fair76, 77 trials evaluated the safety and efficacy of etanercept (25 mg/twice weekly) for the treatment of AS. Studies lasted from 12 to 24 weeks. Overall, these trials included 401 patients. All studies were conducted in patients with moderate to severe AS and allowed concomitant treatment with DMARDs and corticosteroids; one study, however, limited DMARDS to MTX or sulfasalazine.78 Results of all three trials reported that significantly more patients receiving etanercept than placebo presented clinical improvements on all outcome measures (ASAS20/50/70, BASFI [Bath Ankylosing Spondylitis Functional Index], BASDAI) at study endpoint. Significant differences in efficacy started as early as in week 2. Concomitant DMARD treatment did not influence the magnitude of the treatment effect. In the good-rated trial, 57 percent of patients on etanercept and 22 percent of patients on placebo achieved an ASAS20 response after 24 weeks (P < 0.001).78 Patients receiving etanercept also achieved significantly greater positive responses on the majority of secondary outcomes.
Infliximab
Two fair trials assessed the efficacy and safety of infliximab (5 mg/kg) for the treatment of AS.\textsuperscript{79, 80} The larger trial lasted 24 weeks and enrolled 279 patients with moderate to severe AS,\textsuperscript{80} and the smaller study (n = 70) assessed the efficacy and safety of infliximab in patients with severe AS over 12 weeks.\textsuperscript{79} Neither trial allowed concomitant DMARD or corticosteroid treatments. Intention-to-treat results of both trials report significantly greater improvements of infliximab- than of placebo-treated patients on all primary outcome measures (ASAS20/40, BASDAI). After 24 weeks 61 percent of infliximab- and 19 percent of placebo-treated patients achieved an ASAS20 response ($P < 0.001$); 51 percent and 11 percent respectively reported a 50 percent improvement on BASDAI.\textsuperscript{80} However, in this study the mean disease duration was 5.5 years longer in the placebo group than in the infliximab group (no P-value reported) which might bias the treatment effect. In a 2 year open-label extension hospital admissions for infliximab-treated patients were significantly reduced compared to the 12 months before the start of the trial (10\% vs. 41\%).\textsuperscript{83} This corresponds to a reduction of mean inpatient days from 11.1 days before infliximab treatment to 2.9 days after 2 years of treatment.
Table 8. Summary of Efficacy Trials in Adult Patients with AS

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Comparisons</th>
<th>Primary outcome</th>
<th>Secondary outcomes</th>
<th>Population</th>
<th>Results</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ETANERCEPT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calin et al. 200476</td>
<td>RCT</td>
<td>84</td>
<td>12 weeks</td>
<td>ETA+standard treatment / Placebo+standard treatment</td>
<td>ASAS20</td>
<td>ASAS50/70, Schober’s test</td>
<td>Active, moderate to severe AS; mean disease duration: 12.5 yrs.</td>
<td>Response rates on ASAS20/50/70 were significantly greater for ETA than for placebo</td>
<td>Fair</td>
</tr>
<tr>
<td>Davis et al. 200378</td>
<td>RCT</td>
<td>277</td>
<td>24 weeks</td>
<td>ETA+standard treatment / Placebo+standard treatment</td>
<td>ASAS20</td>
<td>ASAS50/70, BASDAI</td>
<td>Active, moderate to severe AS; mean disease duration: 10.3 yrs.</td>
<td>Response rates on ASAS20/50/70 were significantly greater for ETA than for placebo</td>
<td>Good</td>
</tr>
<tr>
<td>Gorman et al. 200277</td>
<td>RCT</td>
<td>40</td>
<td>16 weeks</td>
<td>ETA+standard treatment / Placebo+standard treatment</td>
<td>ASAS20</td>
<td>ASAS50/70, BASFI, Schober’s test</td>
<td>Active, moderate to severe AS; mean disease duration: 13.5 yrs.</td>
<td>Patients on ETA had significantly greater improvements on BASFI and ASAS20 than patients on placebo</td>
<td>Fair</td>
</tr>
<tr>
<td><strong>INFLIXIMAB</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Braun et al. 2002,79</td>
<td>RCT</td>
<td>70</td>
<td>12 weeks</td>
<td>INF / Placebo</td>
<td>BASDAI</td>
<td>BASFI, BASMI, SF-36</td>
<td>Active, moderate to severe AS; mean disease duration: 15.6 yrs.</td>
<td>Patients on INF had significantly greater improvements on BASDAI, BASFI, and SF-36 than patients on placebo</td>
<td>Fair</td>
</tr>
<tr>
<td>Van der Heijde et al. 200580</td>
<td>RCT</td>
<td>279</td>
<td>24 weeks</td>
<td>INF / Placebo</td>
<td>ASAS20</td>
<td>ASAS40, BASDAI</td>
<td>Active, severe AS; mean disease duration: 10.5 yrs.</td>
<td>INF patients had significantly greater improvements on BASDAI, BASFI, and ASAS40 than placebo patients</td>
<td>Fair</td>
</tr>
</tbody>
</table>

AS: ankylosing spondylitis; ASAS: Assessment in Ankylosing Spondylitis; BASDAI: Bath AS Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; BASMI: Bath Ankylosing Spondylitis Metrology Index; ETA: etanercept; INF: infliximab; RCT: randomized controlled trial
IV. Psoriatic arthritis (PsA)
The following drugs are currently approved by the FDA for the treatment of PsA: adalimumab, etanercept, and infliximab.

A. Summary of the Evidence
Overall, the evidence on the comparative effectiveness of TIMs for the treatment of PsA is poor. Fair evidence from two RCTs exists that etanercept is significantly more efficacious than placebo for the treatment of PsA. Two RCTs provide fair evidence on the general efficacy of infliximab and one RCT provides fair evidence that adalimumab is more effective than placebo. Fair evidence from one phase II study exists that alefacept combined with methotrexate is more efficacious than methotrexate alone. Treatment effects are large and consistent across studies. However, significant differences in study characteristics make this evidence insufficient to identify differences in efficacy among TIMs.

B. Description of Studies
For PsA, we did not find any head-to-head trials comparing one TIM to another. We found six placebo-controlled trials assessing the efficacy of etanercept, infliximab, adalimumab and alefacept. The studies ranged in duration from 12 to 50 weeks. We did not find any studies on abetacept, anakinra, efalizumab and rituximab. Included studies are presented in Table 9.

C. Study Populations
All patients suffered from active PsA. However, the definition of active disease varied across studies. Three trials enrolled patients with at least three swollen and three tender joints at screening; two other studies included patients with at least five swollen and five tender joints, and the third study employed additional criteria which utilized clinical sub-types of PsA to establish the presence of PsA. All five trials consisted of patients who had previously failed DMARD and/or MTX therapies.

D. Outcome Measures
All trials assessed response rates as defined by the ACR. In addition, all six studies used the disease specific Psoriatic Arthritis Response Criteria (PsARC) which is composed of a patient global self-assessment, a physician global assessment, a swollen joint score, and a tender joint score. Further details of this scale are presented in Appendix D. In addition, the Psoriasis Area and Severity Index
(PASI) was used in all five studies to measure improvements in both the amount of psoriatic plaque, as well as the severity of the disease. The SF-36 and HAQ were used to assess quality of life. Additionally, one study used a modified Sharp score to assess disease progression.\textsuperscript{88}

**E. Methodological Quality**
All six studies received a fair quality rating. However, the “fair” rating was probably more attributable to poor reporting of methods than to methodological flaws.

**F. Sponsorship**
All trials were funded by the pharmaceutical industry.

**G. Comparative Efficacy and Effectiveness**
We did not identify any head-to-head trials.

**H. General Efficacy**
Because of the lack of head-to-head trials, we reviewed placebo-controlled trials. We have summarized evidence on the general efficacy of TIMs in the treatment of PsA. This, however, does not provide evidence on the comparative efficacy and tolerability of TIMs.

**Adalimumab**
At this time only one trial has been reported on in the literature on the use of adalimumab in PsA.\textsuperscript{93} The included 313 patients suffering from moderate to severe PsA, which was defined as having at least 3 swollen joints and 3 tender or painful joints, who had an inadequate response or intolerance to NSAID therapy. Patients were allowed to continue current methotrexate therapy as long as the dose had been stable for 4 weeks. The double-blinded phase of the study was 24 weeks, but patients who failed to achieve at least a 20 percent decrease in both swollen and tender joint counts on two consecutive visits could receive rescue therapy with corticosteroids or DMARDs. The dose was 40 mg/kg every other week. The adalimumab group saw significantly greater response rates on ACR 20/50/70 than the placebo group (all $P < 0.001$). Sixty percent of the adalimumab group responded according to the PsARC compared to 23 percent on placebo ($P = \text{NR}$).
**Alefacept**

One phase II trial has been reported on in the literature on the use of alefacept in PsA. The study included 185 patients suffering from moderate to severe PsA, which was defined as having at least 3 swollen joints and 3 tender or painful joints, who had an inadequate response to MTX therapy. Patients continued current methotrexate therapy and the dose had been stable for 4 weeks. The double-blinded phase of the study was 24 weeks, 12 weeks of treatment followed by 12 weeks of observation during which methotrexate treatment was continued in all participants. The dose was 15 mg every week. The adalimumab group saw significantly greater response rates on ACR 20 than the placebo group, 54 percent versus 23 percent ($P < 0.001$). There were no significant differences in the other outcomes which included ACR 50/70, PASI and PGA (physician global assessment), though there was a trend that favored alefacept, ACR50/70 was achieved by 17 percent and 7 percent of the alefacept group versus 10 percent and 2 percent, respectively, of the placebo group. Similarly the PASI 50 and a PGA of clear or almost clear was reported in 45 percent and 31 percent of the alefacept group versus 31 percent and 24 percent in the placebo group.

**Etanercept**

Two fair studies examined the efficacy of etanercept in patients with PsA. Overall, 265 patients with active PsA, not adequately responding to standard DMARD therapies, were included. In both studies patients were allowed to continue MTX therapy as long as it had been stable for four weeks prior. One study lasted 12 weeks; the other trial was double-blinded for 24 weeks. Both studies had the same dosing regimen of 25 mg of etanercept twice-weekly subcutaneous injections. In both studies response rates compared to placebo on ACR20 were significantly greater for etanercept. In the 12 week study, 87 percent of the patients on etanercept achieved a PsARC response compared to 23 percent on placebo ($P < 0.0001$). The longer study had similar results in patients achieving a PsARC response at 12 weeks; 72 percent of the patients on etanercept responded versus 31 percent on placebo. Quality of life was significantly improved as measured by the HAQ in both studies. Mean improvements were 83 percent in etanercept- compared to 3 percent in placebo-treated patients in the 12 week study ($P < 0.0001$). In the longer study, at 24 weeks the mean improvement was 54 percent in the etanercept group and 6 percent in the placebo group ($P < 0.0001$). The longer study assessed the radiographic progression of disease at 24 weeks and found the annualized modified Sharp score was significantly less in etanercept- than in placebo-treated patients ($P = 0.0001$).
Infliximab
We found two fair studies on the use of infliximab in patients with PsA.\textsuperscript{89-92} Overall, 304 patients with active PsA, not adequately responding to standard DMARD therapies, were included. In both studies patients were allowed to continue MTX therapy as long as it had been stable for four weeks prior. The earlier study was double-blinded for 16 weeks;\textsuperscript{89} the other trial was double-blinded for 24 weeks with cross-over allowed at week 16 for non-responders.\textsuperscript{90} Both studies had the same dosing regimen of 5 mg/kg of infliximab at weeks 0, 2, 6, 14 and the longer study had an additional injection at week 22. In both studies response rates compared to placebo on ACR20 were significantly greater for infliximab. In the earlier study, 86 percent of the patients on infliximab achieved a PsARC response compared to 12 percent on placebo ($P < 0.001$).\textsuperscript{89} The bigger study had similar results in patients achieving a PsARC response at 14 weeks; 77 percent of the patients on infliximab responded versus 27 percent on placebo.\textsuperscript{90} Quality of life was significantly improved as measured by the HAQ in both studies. Mean improvements were 49.8 percent in infliximab compared to -1.6 percent in placebo-treated patients in the smaller study ($P < 0.001$). In the bigger study, at 14 weeks the mean improvement was 48.6 percent in the infliximab group and an 18.4 percent loss in the placebo group ($P < 0.001$).
Table 9. Summary of Efficacy Trials in Adult Patients with PsA

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Comparisons</th>
<th>Primary outcome</th>
<th>Secondary outcomes</th>
<th>Population</th>
<th>Results</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ADALIMUMAB</strong></td>
<td>RCT</td>
<td>313</td>
<td>24 weeks</td>
<td>ADA + MTX / Placebo + MTX</td>
<td>ACR 20, change in modified Sharps score</td>
<td>ACR50/70, HAQ, PsARC, SF-36</td>
<td>Active PsA; failed at least one DMARD; mean disease duration: 9.5 years</td>
<td>ADA had significantly better outcomes than placebo</td>
<td>Fair</td>
</tr>
<tr>
<td>Mease et al. 2005</td>
<td>RCT</td>
<td>185</td>
<td>24 weeks (12 weeks treatment, 12 weeks observation)</td>
<td>ALE + MTX / Placebo + MTX</td>
<td>ACR 20</td>
<td>Active PsA; failed at least one DMARD; mean disease duration: NR</td>
<td>ALE had significantly better ACR20 than placebo</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td><strong>ETANERCEPT</strong></td>
<td>RCT</td>
<td>60</td>
<td>12 weeks</td>
<td>ETA + MTX / Placebo + MTX</td>
<td>PsARC, PASI</td>
<td>ACR20/50/70, HAQ</td>
<td>Active PsA; failed at least one DMARD; median disease duration: 10 years</td>
<td>ETA had significantly better outcomes than placebo</td>
<td>Fair</td>
</tr>
<tr>
<td>Mease et al. 2000</td>
<td>RCT</td>
<td>205</td>
<td>72 weeks (24 blinded, 48 open-label)</td>
<td>ETA + MTX / MTX + Placebo</td>
<td>ACR 20</td>
<td>Active PsA; failed at least one DMARD; mean disease duration 9.1 years</td>
<td>ETA had significantly better outcomes than placebo</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td><strong>INFLIXIMAB</strong></td>
<td>RCT</td>
<td>104</td>
<td>50 weeks</td>
<td>INF/ Placebo (71% received a concomitant DMARD)</td>
<td>ACR20 and PASI</td>
<td>ACR 50/70 DAS; HAQ; ratings of enthesitis and dactylitis; PsARC.</td>
<td>Active PsA; failed at least one DMARD; mean disease duration 11.4 years</td>
<td>INF had significantly better outcomes than placebo</td>
<td>Fair</td>
</tr>
<tr>
<td>Antoni et al. IMPACT Study 2005</td>
<td>RCT</td>
<td>200</td>
<td>24 weeks</td>
<td>INF/ Placebo (46% received concomitant MTX)</td>
<td>ACR20; HAQ; SF-36</td>
<td>ACR50/70; PsARC; PASI; dactylitis and enthesopathy</td>
<td>Active PsA; failed at least one DMARD; mean disease duration 8 years</td>
<td>INF had significantly better outcomes than placebo</td>
<td>Fair</td>
</tr>
</tbody>
</table>

ACR: American College of Rheumatology; ADA: adalimumab; ALE: alfacept; DAS: disease activity score; DMARD: disease-modifying antirheumatic drug; ETA: etanercept; HAQ: Health Assessment Questionnaire; INF: infliximab; MTX: methotrexate; NR: not reported; PsARC: psoriatic arthritis response criteria; SF-36: Medical Outcomes Study Short Form 36 Health Survey
V. Crohn’s Disease

Only infliximab currently is approved by the FDA for the treatment of Crohn’s disease.

A. Summary of the evidence

Overall, the evidence on the comparative effectiveness of TIMs for the treatment Crohn’s Disease is poor. No evidence directly comparing the efficacy and safety of one TIM to another could be found, and evidence was insufficient to make indirect comparisons.

A single study rated as good suggests that adalimumab is more efficacious than placebo for moderate to severe Crohn’s disease. However, this study was only 4 weeks in duration. Fair to good evidence from six RCTs exists that infliximab is significantly more efficacious than placebo for initial (i.e., patients with refractory Crohn’s disease that had not received a TIM during the previous 12 weeks) and maintenance treatment of Crohn’s disease. Treatment effects are large and evident within 1 to 2 weeks. On average, a two to three-fold increase in the number of responders was observed among infliximab-treated patients compared to placebo. Maintenance treatment with infliximab maintains a response significantly longer than placebo, although infections and infusion-related reactions are more common with long-term treatment. Infliximab is also more efficacious than placebo in fistulizing Crohn's disease (a serious complication of Crohn’s disease characterized by abnormal communication between the gut and the skin, with small bowel or colonic contents draining to the skin surface). Fair evidence from one small RCT exists that etanercept is no more efficacious than placebo and adverse reactions are more common in etanercept- than placebo-treated patients. We did not find any evidence on the efficacy and safety of abatacept, alefacept, anakinra, efalizumab, and rituximab for the treatment of Crohn’s disease.

Although some studies allowed stable doses of other immunomodulatory agents, no conclusive evidence exists to determine whether combination treatment of etanercept and infliximab with other agents (azathioprine, 6-MP, MTX) leads to clinically and statistically greater improvements than monotherapy.

B. Description of Studies

For Crohn’s disease, we did not find any head-to-head RCTs comparing one TIM to another. We found six placebo-controlled trials and two observational studies that assessed the efficacy and safety of infliximab. We also identified one trial that compared the efficacy and safety of adalimumab to placebo.
and one trial that compared etanercept to placebo. We did not find any studies on abatacept, alefacept, anakinra, efalizumab, or rituximab. Included studies are presented in Table 10.

C. Study Populations
All patients suffered from active Crohn’s disease of at least 3 months’ duration. Some patients also had abdominal or perianal fistulas. Most studies included patients with a Crohn’s Disease Activity Index (CDAI) between 220 and 400. However, some trials included patients with CDAI scores as high as 450 (i.e., more severe disease). The non-randomized studies were population-based and followed consecutive patients treated with infliximab. One study included patients with other inflammatory bowel diseases, including ulcerative colitis and indeterminate colitis; however, 88 percent of patients had a diagnosis of Crohn’s disease. Disease duration and concomitant treatments varied across studies. On average, disease duration ranged from 8 to 12 years. Many studies allowed concomitant treatment with 5-ASA, antibiotics, corticosteroids, azathioprine, 6-MP, or MTX.

D. Outcome Measures
Most studies utilized the National Cooperative Crohn’s Disease Study rating scale, the CDAI, to characterize disease severity. The CDAI assesses eight related variables (e.g., number of liquid or soft stools per day, severity of abdominal pain or cramping, general well-being, the presence or absence of extraintestinal manifestations of disease, the presence or absence of abdominal mass, the use or nonuse of antidiarrheal drugs, the hematocrit, and body weight; see Appendix D) to yield a composite score between 0 and 600; scores below 150 indicate remission while scores above 450 indicate severe illness. Response commonly was characterized by a CDAI reduction greater than or equal to 70 points. Several studies utilized the Inflammatory Bowel Disease Questionnaire (IBDQ). The IBDQ identifies 32 individual items categorized within four major quality of life domains (primary bowel symptoms, systemic symptoms, social impairment, and altered emotional function). Some studies assessed CRP concentrations as an intermediate marker for inflammation. In studies specifically designed to assess fistulizing disease, outcomes included 50 percent reduction in the number of draining fistulas or a complete absence in draining fistulas.

E. Methodological Quality
One trial was rated as “good” quality. Although all other included trials were given a “fair” quality rating, study quality varied. Several trials did not report the number of patients lost to follow up, and some trials had loss to follow-up exceeding 50%. Smaller trials may not have had sufficient sample size to detect differences in health outcomes (from a patient’s perspective). Randomization methods and
blinding were generally adequate; all studies used a double-dummy design (i.e., using 0.1% human serum albumin placebo in an identical container to active treatment) to guarantee blinding; method of allocation concealment was rarely reported.

F. Sponsorship
All studies, except the observational studies, were funded by the pharmaceutical industry. Several studies also received funding from the National Institutes of Health or the FDA.

G. Comparative Efficacy and Effectiveness
We did not identify any head-to-head RCTs or observational studies. Additionally, we were unable to make indirect comparisons because there were too few trials and existing trials were too different in design.

H. General Efficacy
Because of the lack of head-to-head trials, we reviewed placebo-controlled trials. We summarized evidence on the general efficacy of TIMs in the treatment of Crohn’s disease; however, this does not provide evidence on the comparative efficacy and tolerability of TIMs.

Adalimumab
A single trial rated good compared adalimumab to placebo.\textsuperscript{95} Two hundred and ninety-nine patients with moderate to severe Crohn’s disease (CDAI score 220 to 450) were randomized to receive two doses of subcutaneous placebo or adalimumab. Adalimumab was given as three different induction regimens: 40mg at week 0 and 20mg at week 2; 80mg at week 0 and 40mg at week 2; 160mg at week 0 and 80mg at week 2. Primary analyses compared CDAI remission rates (<150 points) at week 4 for the 160/80mg and 80/40mg dosing arms to placebo. Secondary analyses assessed CDAI response (> 70 point or > 100 point change) and IBDQ scores. Groups were well balanced at baseline. At 4 weeks, both the 160/80mg and 80/40mg dosing arms had statistically significantly more CDAI remitters than placebo (36% and 24% compared to 12%, respectively; \(P = 0.001\)). CDAI response assessed with a 70 point reduction also was significantly better than placebo for the 160/80mg and 80/40mg groups (\(P<0.01\), although only the 160/80mg group was significantly better than placebo for the 100 point reduction (\(P=0.002\)). Adverse events occurred at similar frequencies for placebo and adalimumab treated patients, except for injection site reactions which were more common among adalimumab patients.
**Etanercept**

A single fair trial compared etanercept to placebo.\(^98\) Forty-three patients with moderate to severe Crohn’s disease (CDAI score 220 to 450) were randomized to receive subcutaneous placebo or etanercept 25 mg twice weekly for 8 weeks. Patients were at least 12 years of age and could not have taken another TIM within 12 weeks. Primary outcome measures were clinical response (CDAI decrease ≥ 70 points) or remission (CDAI score < 150). No statistically significant differences between etanercept and placebo in clinical response or remission were detected at any time. Furthermore, no differences in quality of life or the rate of fistula improvement were observed. Compared to placebo, more etanercept-treated patients reported adverse events (74% vs. 50%; P-value not reported); injection site reactions and headache were the most commonly reported adverse events.

**Infliximab**

Six fair trials compared infliximab to placebo.\(^99\)–\(^104\) Two trials assessed the efficacy of a single infliximab infusion,\(^99\)\(^,\)\(^104\) and two trials assessed the efficacy of repeated maintenance infusions.\(^100\),\(^102\) Two additional trials compared infliximab to placebo in patients with Crohn’s disease with multiple draining abdominal or perianal fistulas.\(^101\),\(^103\) Two uncontrolled studies reported the efficacy and tolerability of infliximab in consecutively treated patients with inflammatory bowel disease (including Crohn’s disease, ulcerative colitis, and indeterminate colitis).\(^96\),\(^97\)

Two trials examined the efficacy of a single infusion of infliximab at doses of 5, 10, and 20 mg/kg in Crohn’s disease (CDAI scores between 220 and 400).\(^99\),\(^104\) Randomized patients were refractory to corticosteroids, mesalamine, 6-mercaptopurine, or azathioprine. Both trials demonstrated significantly better efficacy of a single infusion of infliximab compared to placebo. In the smaller European trial, 30 patients with active Crohn’s disease were randomized to a single 5, 10, or 20 mg/kg dose of infliximab or placebo.\(^99\) At 4 weeks, all patients underwent a full colonoscopy and ileoscopy and a Crohn’s Disease Endoscopy Index of Severity (CDEIS) score was calculated; CDAI scores and CRP concentrations also were assessed. All doses of infliximab were significantly better than placebo at 4 weeks (\(P < 0.05\)). In the 12 week multinational trial,\(^104\) 108 patients randomized to infliximab 5, 10, or 20 mg/kg or placebo were assessed at 2, 4, and 12 weeks. Responders were characterized as having a CDAI reduction of 70 points or more. Quality of life with respect to bowel function (IBDQ) and CRP concentrations also were assessed. At 4 weeks, compared to placebo, significantly more infliximab-treated patients were characterized as CDAI responders (\(P < 0.005\)). Quality of life scores and CRP concentrations also were significantly better than placebo in patients treated with infliximab (\(P < 0.05\) and \(P < 0.01\), respectively).\(^105\)
To assess the ability of infliximab to maintain treatment response, maintenance infusions of infliximab were compared to placebo in a 36 week and a 54 week trial.\textsuperscript{100,102} In both trials, patients with Crohn’s disease (CDAI scores between 220 and 400) responding to an initial infliximab infusion were randomized. One trial was a continuation of the 12 week trial described above;\textsuperscript{104} in this trial 73 patients responding to the initial 5, 10, or 20 mg/kg infusion of infliximab were randomized to receive infliximab 10 mg/kg repeated at 8-week intervals for four additional doses or placebo.\textsuperscript{102} Retreatment with infliximab maintained the initial treatment benefit in 62\% of patients compared to 37\% of placebo-treated patients ($P = 0.16$). In the ACCENT 1 trial,\textsuperscript{100} 335 patients responding (CDAI decrease $\geq 70$ points) at 2 weeks to an initial infliximab infusion of 5 mg/kg were randomized to repeat infusions of placebo, infliximab 5 mg/kg, or infliximab 10 mg/kg at week 2 and 6 and then every 8 weeks thereafter until week 46. Primary outcome measures included time to loss of response (CDAI $\geq 175$) and the proportion of week 2 responders in remission (CDAI < 150) at week 30. Compared to placebo, infliximab-treated patients had a significantly longer time to loss of response ($P < 0.001$) and the odds of being in remission at week 30 were nearly three times greater. Infliximab maintenance therapy demonstrated greater mucosal healing compared with the placebo maintenance group at both weeks 10 and 54. Infliximab-treated patients also had better endoscopic healing, fewer hospitalizations, fewer surgeries, decreased corticosteroid use, fewer hours lost from work, and better quality of life scores ($P < 0.05$ for all).\textsuperscript{106-108} Additional analyses found scheduled maintenance treatment with infliximab to have better mucosal healing than episodic treatment ($P=0.007$).\textsuperscript{109}

Two trials\textsuperscript{101,103} compared the efficacy and safety of infliximab to placebo in patients with enterocutaneous or perianal fistulas, a serious complication of Crohn’s disease characterized by abnormal communication between the gut and the skin with small bowel or colonic contents draining to the skin surface.\textsuperscript{101,103} A 34 week study randomized 94 adult patients who had abdominal or perianal fistulas of at least 3 months’ duration as a complication of Crohn’s disease to placebo, 5 mg/kg infliximab, or 10 mg/kg infliximab.\textsuperscript{101} Doses were administered intravenously at baseline, 2 and 6 weeks. Compared to placebo, significantly more infliximab-treated patients had a reduction of 50\% or more from baseline in the number of draining fistulas observed at 2 or more consecutive visits ($P < 0.05$). Likewise, 55 percent of patients on infliximab 5 mg/kg and 38 percent of patients on 10 mg/kg had closure of all fistulas, compared to 13 percent of patients assigned to placebo ($P = 0.001$ and $P = 0.04$, respectively). In the ACCENT II trial,\textsuperscript{103} 195 patients with Crohn’s disease and one or more draining abdominal or perianal fistulas who responded to 3 open-label 5 mg/kg infusions of infliximab were randomized to maintenance treatment with 8-week infusions of infliximab 5 mg/kg or placebo. Patients that did not respond to open-label treatment (n = 87) also were followed for safety. The primary outcome was defined as time to loss
Observational evidence of efficacy comes from two case series studies.\textsuperscript{96, 97} A Stockholm County, Sweden, population based cohort study supports the general efficacy of infliximab in patients with inflammatory bowel disease.\textsuperscript{96} Among 217 consecutive patients treated with infliximab (191 patients had Crohn’s disease), 75 percent (n = 163) demonstrated at least some degree of response; 48 percent of patients (n = 104) achieved remission. However, a 2.8 percent mortality rate was observed, emphasizing the need for vigilance in drug surveillance. A second case series analysis in Edmonton, Alberta, reviewed 109 consecutive patients with inflammatory and/or fistulizing Crohn’s disease who received infliximab.\textsuperscript{97} A clinical response was documented in 73 percent (n = 80) of patients; 55 percent of patients (n = 61) had a partial response and 17 percent (n = 19) had a full response. No deaths were reported.

Of response. On average, patients randomized to infliximab maintenance therapy maintained their response for more than 26 weeks longer than placebo ($P < 0.001$). At week 54, 36 percent of infliximab-treated patients had a complete absence of draining fistulas compared to 19% of placebo-treated patients ($P = 0.009$). At 6 weeks, infliximab also was more efficacious than placebo in a subgroup of women with rectovaginal fistulas (fistula closure 61% and 45%, respectively).\textsuperscript{110} Compared to placebo, infliximab-treated patients had fewer hospitalizations (11 vs. 31; $P < 0.05$), fewer mean hospitalization days (0.5 vs. 2.5 days/100; $P < 0.05$), and fewer surgeries and procedures (65 vs. 126; $P < 0.05$).\textsuperscript{111} No differences between active treatment and placebo were found in the number of fistula-related abscesses.\textsuperscript{112}
### Table 10. Summary of Efficacy Trials in Adult Patients with Crohn’s Disease

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Comparisons</th>
<th>Primary outcome</th>
<th>Secondary outcomes</th>
<th>Population</th>
<th>Results</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hanauer et al., 2006&lt;sup&gt;95&lt;/sup&gt;</td>
<td>RCT</td>
<td>299</td>
<td>4 weeks</td>
<td>ADA / placebo</td>
<td>CDAI remission (&lt;150)</td>
<td>IBDQ; CDAI response (&lt; 70 or &lt; 100)</td>
<td>Patients 18-75 years with moderate to severe Crohn’s disease</td>
<td>Significantly more remission and better IBDQ scores for 80/40mg dose and 160/80mg dose than for placebo</td>
<td>Good</td>
</tr>
<tr>
<td>Sandborn et al., 2001&lt;sup&gt;98&lt;/sup&gt;</td>
<td>RCT</td>
<td>43</td>
<td>8 weeks</td>
<td>ETA / placebo</td>
<td>CDAI</td>
<td>Rate of fistula improvement, fistula closure, IBDQ</td>
<td>Patients 12 and older with moderate to severe Crohn’s disease</td>
<td>No difference between ETA and placebo in response, remission, quality of life, or fistula improvement</td>
<td>Fair</td>
</tr>
<tr>
<td>D’Haens et al., 1999&lt;sup&gt;99&lt;/sup&gt;</td>
<td>RCT</td>
<td>30</td>
<td>4 weeks</td>
<td>INF / placebo</td>
<td>CDEIS, CRP</td>
<td>&gt; 6 month history of moderate to severe active Crohn’s disease refractory to corticosteroids, mesalamine, 6-mercaptopurine, or azathioprine</td>
<td>INF-treated patients more likely to sustain clinical response, had a longer time to loss of response, better quality of life, better endoscopic healing, fewer surgeries and hospitalizations, and less work loss than placebo-treated patients</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Hanauer et al., 2002&lt;sup&gt;100&lt;/sup&gt;, 106-109</td>
<td>RCT</td>
<td>573</td>
<td>54 weeks</td>
<td>INF / placebo</td>
<td>Proportion of week 2 responders in remission at week 30; time to loss of response</td>
<td>Employment status/work loss, surgeries, SF-36, IBDQ, hospitalizations, corticosteroid discontinuation, endoscopic healing</td>
<td>&gt; 3 month history of moderate to severe Crohn’s disease and CDAI response at 2 weeks to single dose 5mg/kg INF</td>
<td>INF-treated patients more likely to sustain clinical response, had a longer time to loss of response, better quality of life, better endoscopic healing, fewer surgeries and hospitalizations, and less work loss than placebo-treated patients</td>
<td>Fair</td>
</tr>
</tbody>
</table>
Table 10: Summary of Efficacy Trials in Adult Patients with Crohn’s Disease (continued)

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Comparisons</th>
<th>Primary outcome</th>
<th>Secondary outcomes</th>
<th>Population</th>
<th>Results</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ljung et al., 2004&lt;sup&gt;66&lt;/sup&gt;</td>
<td>Case series</td>
<td>217</td>
<td></td>
<td>INF</td>
<td>Adverse events</td>
<td>Clinical response, remission, failure</td>
<td>Consecutive patients with IBD treated with infliximab between January 1999- April 2001</td>
<td>Overall response rate was 75% with 48% of patients achieving remission</td>
<td>N/A</td>
</tr>
<tr>
<td>Present et al., 1999&lt;sup&gt;101&lt;/sup&gt;</td>
<td>RCT</td>
<td>94</td>
<td>34 weeks</td>
<td>INF / placebo</td>
<td>Reduction of 50% or more in the number of draining fistulas</td>
<td>Closure of all fistulas, time to beginning of response and duration of response, CDAI, PDAI</td>
<td>Adults with Crohn’s disease with multiple draining abdominal or perianal fistulas of at least 3 months’ duration</td>
<td>Significantly greater reduction in the number of draining fistulas, shorter time to response, and greater improvement in PDAI for INF compared to placebo; no difference in CDAI at endpoint</td>
<td>Fair</td>
</tr>
<tr>
<td>Rutgeerts et al., 1999&lt;sup&gt;102&lt;/sup&gt;</td>
<td>RCT</td>
<td>73</td>
<td>36 weeks</td>
<td>INF / placebo</td>
<td>Maintained response (CDAI ≥ 70) or remission (CDAI &lt; 150), discontinuation rate (efficacy)</td>
<td>Mean CDAI, IBDQ, CRP</td>
<td>&gt; 6 months history of moderate to severe active Crohn’s disease and previous response to INF</td>
<td>Statistically modest improvements in response, remission, time to loss of response, CDAI, IBDQ and CRP for INF compared to placebo</td>
<td>Fair</td>
</tr>
<tr>
<td>Sample et al., 2002&lt;sup&gt;97&lt;/sup&gt;</td>
<td>Case series</td>
<td>109</td>
<td>≥ 8 weeks</td>
<td>INF</td>
<td>Adverse events</td>
<td>Clinical response, corticosteroid tapering</td>
<td>Consecutive patients with Crohn’s disease treated with INF</td>
<td>73% of INF-treated patients had a clinical response and steroids were tapered in 53%; AEs 7%</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**INFLIXIMAB**
Table 10: Summary of Efficacy Trials in Adult Patients with Crohn’s Disease (continued)

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Comparisons</th>
<th>Primary outcome</th>
<th>Secondary outcomes</th>
<th>Population</th>
<th>Results</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sands et al., 2004&lt;sup&gt;103,110-112&lt;/sup&gt;</td>
<td>RCT</td>
<td>282</td>
<td>54 weeks</td>
<td>INF / placebo</td>
<td>Time to loss of response after randomization (week 14)</td>
<td>CDAI, IBDQ, hospitalizations, hospitalization days, surgeries</td>
<td>&gt; 3 month history of active Crohn’s with multiple draining fistulas and 14 week response (≥ 50% closure) to 3 open label doses of INF 5mg/kg</td>
<td>Significantly longer time to loss of response, fewer draining fistulas, greater improvement in CDAI and IBDQ, fewer hospitalizations, hospitalization days, and surgeries for INF compared to placebo; no difference in fistula-related abscesses for maintenance</td>
<td>Good</td>
</tr>
<tr>
<td>Targan et al., 1997&lt;sup&gt;104,105&lt;/sup&gt;</td>
<td>RCT</td>
<td>108</td>
<td>12 weeks</td>
<td>INF / placebo</td>
<td>Response at 4 weeks (≥ 70 point reduction in CDAI)</td>
<td>IBDQ, CRP</td>
<td>&gt; 6 month history of moderate to severe Crohn’s disease refractory to corticosteroids, mesalamine, 6-mercaptopurine, or azathioprine</td>
<td>Significantly more responders and greater improvement in IBDQ and CRP for INF compared to placebo</td>
<td>Fair</td>
</tr>
</tbody>
</table>

CDAI: Crohn’s Disease Activity Index; CRP: C-reactive protein; ETA: etanercept; IBDQ: Inflammatory-bowel-disease questionnaire; INF: infliximab; PDAI: Pouchitis Disease Activity Index
VI. Ulcerative colitis (UC)
Infliximab is the only drug currently approved by the FDA for the treatment of UC.

A. Summary of the Evidence
Overall, the evidence on the comparative effectiveness of TIMs for the treatment of UC is poor. Fair evidence from three RCTs exists that infliximab is significantly more efficacious than placebo for the treatment of active UC. Treatment effects are large across studies.

B. Description of Studies
For UC, we did not find any head-to-head trials comparing one TIM to another. We found three placebo-controlled trials (2 publications) assessing the efficacy of infliximab.\textsuperscript{113,114} The studies ranged in duration from 12 to 54 weeks. We did not find any studies on abetacept, adalimumab, alefacept, anakinra, efalizumab, etanercept, and rituximab. Included studies are presented in Table 11.

C. Study Populations
All patients suffered from active UC. However, the definition of active disease varied across studies. One trial enrolled patients with an active severe to moderately severe attack of UC based on bowel movements, bloody feces and high CRP levels.\textsuperscript{114} Two fair studies, reported in the same article, included patients with moderate to severe UC based on stool frequency, rectal bleeding, endoscopy and physician’s assessment.\textsuperscript{113} All three trials consisted of patients who had previously failed ASA-5 and steroid treatments.

D. Outcome Measures
One fair study used the primary outcomes of death and colectomy to assess the efficacy of infliximab versus placebo.\textsuperscript{113} The other two fair studies used clinical response, a decrease of 30% in the presenting symptoms at 8 weeks as the primary outcome.\textsuperscript{113} In addition they also captured rates of response and remission at later times—30 and 54 weeks.\textsuperscript{113}
E. Methodological Quality
All three trials were rated fair, although two of these trials suffered from high rates of attrition, particularly in the placebo arms.

F. Sponsorship
All trials were funded by the pharmaceutical industry.

G. Comparative Efficacy and Effectiveness
We did not identify any head-to-head trials.

H. General Efficacy
Because of the lack of head-to-head trials, we reviewed placebo-controlled trials. We have summarized evidence on the general efficacy of TIMs in the treatment of UC. This, however, does not provide evidence on the comparative efficacy and tolerability of TIMs.

Infliximab
We found three fair studies (2 publications) on the use of infliximab in patients with UC.\textsuperscript{113, 114} The two larger studies were double-blinded for 30 or 54 weeks and each consisted of 364 patients.\textsuperscript{113} These two studies, termed ACT 1 (Active Ulcerative Colitis Trial) and ACT 2, had dosing regimens of 5 or 10 mg/kg at weeks 0, 2, 6 and then every eight weeks. Concomitant medications were continued except for corticosteroids which were tapered down by 5 mg per week until a dose of 20 mg was reached and then additional reductions occurred at a rate of 2.5 mg per week. ACT 1 and ACT 2 showed clinical responses at 8 weeks that were significantly better in the infliximab groups. In ACT 1, at 8 weeks, 69 percent of patients receiving 5 mg/kg and 62 percent receiving 10 mg/kg responded versus 37 percent placebo patients (for both $P < 0.001$). Similarly in ACT 2, at eight weeks 65 percent patients receiving 5 mg/kg and 69 percent receiving 10 mg/kg responded versus 29 percent placebo patients (for both, $P < 0.001$), however the attrition rates were very high at the study endpoints of 30 and 54 weeks and not reported at eight weeks when the primary outcome was evaluated. ACT 1 had attrition of 37 percent in patients receiving 5 mg/kg and 40 percent receiving 10 mg/kg responded versus 61 percent placebo patients and ACT 2 had attrition of 19 percent in patients receiving 5 mg/kg and 22 percent receiving 10 mg/kg responded versus 46 percent placebo patients.

In the third study, 45 patients with active UC, not adequately responding to standard corticosteroid therapy, were included.\textsuperscript{114} All patients continued to receive 5-ASA treatment based on the individual investigator’s opinion. The study was double blinded for 12 weeks and patients received one 5 mg/kg dose of infliximab or placebo. Results showed a reduction in colectomy in the group receiving infliximab.
at 29 percent versus placebo, in which 67 percent underwent colectomy within three months ($P = 0.017$). The corresponding odds ratio was 4.9 with a 95% confidence interval ranging 1.4 to 17.
Table 11. Summary of Efficacy Trials in Adult Patients with Ulcerative Colitis

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Comparisons</th>
<th>Primary outcome</th>
<th>Secondary outcomes</th>
<th>Population</th>
<th>Results</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarnerot et al. 2005&lt;sup&gt;14&lt;/sup&gt;</td>
<td>RCT</td>
<td>45</td>
<td>12 weeks</td>
<td>INF vs. Placebo</td>
<td>Colectomy or death</td>
<td>ACR50/70, HAQ, PsARC, SF-36</td>
<td>Acute severe or moderate UC attack; nonresponsive to IIVT or steroids; mean disease duration: NR</td>
<td>INF had significantly better colectomy outcomes than placebo</td>
<td>Fair</td>
</tr>
<tr>
<td>Rutgeerts et al. 2005&lt;sup&gt;13&lt;/sup&gt; Act I</td>
<td>RCT</td>
<td>364</td>
<td>54 weeks</td>
<td>INF vs. Placebo</td>
<td>Clinical response at week 8</td>
<td>Clinical response at weeks 30 and 54; clinical remission and mucosal healing</td>
<td>Moderate to severe UC; nonresponsive or intolerant to ASA-5 and/or steroids; mean disease duration: 6.8 years</td>
<td>INF had significantly better response rates than placebo</td>
<td>Fair</td>
</tr>
<tr>
<td>Rutgeerts et al. 2005&lt;sup&gt;13&lt;/sup&gt; Act II</td>
<td>RCT</td>
<td>364</td>
<td>30 weeks</td>
<td>INF vs. Placebo</td>
<td>Clinical response at week 8</td>
<td>Clinical response at weeks 30 and 54; clinical remission and mucosal healing</td>
<td>Moderate to severe UC; nonresponsive or intolerant to ASA-5 and/or steroids; mean disease duration: 6.5 years</td>
<td>INF had significantly better response rates than placebo</td>
<td>Fair</td>
</tr>
</tbody>
</table>

ACR: American College of Rheumatology; ASA: Assessment in Ankylosing Spondylitis; HAQ: Health Assessment Questionnaire; INF: infliximab; NR: not reported; PsARC: Psoriatic Arthritis Response Criteria; SF-36: Medical Outcomes Study Short Form 36 Health Survey; UC: ulcerative colitis
VII. Plaque Psoriasis
The following drugs are currently approved by the FDA for the treatment of plaque psoriasis: alefacept, efalizumab, etanercept, and infliximab.

A. Summary of the evidence
Overall, the evidence on the comparative effectiveness of TIMs for the treatment of plaque psoriasis is poor. No evidence directly comparing the efficacy and safety of one TIM to another could be found, and evidence was insufficient to make indirect comparisons. Fair to good evidence exists on the general efficacy of alefacept, etanercept, efalizumab, and infliximab for the treatment of plaque psoriasis.

B. Description of Studies
For plaque psoriasis, we did not find any head-to-head RCTs comparing one TIM to another. We found 13 placebo-controlled trials that assessed the efficacy and safety of TIMs for the treatment of plaque psoriasis: two on alefacept, four on etanercept, four on efalizumab, and three on infliximab. We did not find any studies on abatacept, adalimumab, anakinra, or rituximab. Included studies are presented in Table 12.

C. Study Populations
Studies, in general, enrolled patients who had a history of plaque psoriasis for more than 6 months, with more than 10 percent of body area involved. Minimum PASI scores to meet inclusion criteria ranged from 10 to 12. Most patients had had previous systemic treatments for plaque psoriasis or were candidates for systemic treatments. Patients were excluded if they had clinically significant disease flares at screening or enrollment, major concomitant illnesses, immune disorders, or organ dysfunction. Prior therapy with biologic agents was an exclusion criterion for most studies.

D. Outcome Measures
All studies assessed PASI 50 or PASI 75 as one of the primary outcome measures. In addition, most trials included some measure of quality of life or functional capacity such as the Dermatology Life Quality Index (DQLI), Dermatology Quality of Life Scale (DQOLS), the itching VAS (visual analogue scale), or the SF-36.

E. Methodological Quality
The methodological quality of studies was generally good and some of the “fair” ratings are probably more attributable to inadequate reporting than methodological flaws. Randomization methods and
blinding were generally adequate; all studies used a double-dummy design (i.e., using 0.1% human serum albumin placebo in an identical container to active treatment) to guarantee blinding; method of allocation concealment was rarely reported.

F. Sponsorship
The majority of studies (84 percent) were funded by the pharmaceutical industry. One trial (8 percent) received funding from a governmental agency as well as from the industry. For one study, we could not determine the source of funding.

G. Comparative Efficacy and Effectiveness
We did not identify any head-to-head RCTs or observational studies. Additionally, we were unable to make indirect comparisons because there were too few trials and existing trials were too different in design.

H. General Efficacy
Because of the lack of head-to-head trials, we reviewed placebo-controlled trials. We summarized evidence on the general efficacy of TIMs in the treatment of plaque psoriasis; however, this does not provide evidence on the comparative efficacy and tolerability of TIMs.

Alefacept
One good and one fair study provide evidence on the general efficacy of alefacept for the treatment of plaque psoriasis. Both trials lasted 24 weeks, during the first 12 weeks patients received intramuscular alefacept treatment. The larger trial (n = 507) was a multinational study that randomized patients to placebo, 10 mg or 15 mg (FDA approved dosage) of alefacept administered once weekly for 12 weeks. Throughout the study, the proportion of patients achieving 75 percent PASI reduction from baseline was significantly higher (P < 0.001) in patients receiving 15mg (33%) and 10 mg of alefacept (28%) than in those treated with placebo (13%). Likewise, quality of life outcomes (SF-36, DLQI, DQOLS) were significantly more improved in patients on 15mg of alefacept than on placebo. Patients treated with 10mg of alefacept had greater improvements than those on placebo, however, the differences did not always reach statistical significance. The smaller study (n=229) was a dose-ranging trial that provided similar results with respect to PASI and quality of life outcomes as the good-quality study.

Efalizumab
Three fair and one good trials provide evidence on the efficacy and safety of efalizumab for the treatment of plaque psoriasis. All four studies conducted double-blinded comparisons over 12 weeks and

Targeted Immune Modulators
provide consistent evidence on the general efficacy of efalizumab. Two trials were dose-ranging studies that also used higher than FDA approved dosages (FDA approved dosage: 1mg/kg).\textsuperscript{120,121} The study rated as good quality enrolled 556 patients with moderate to severe plaque psoriasis.\textsuperscript{123} Patients were randomized to efalizumab (1mg/kg) or placebo. After 12 weeks 26.6 percent of patients on efalizumab and 4.3 percent on placebo achieved a PASI 75 response. The efalizumab group also had statistically significantly better outcomes on PGA and DLQI. Similar treatment effects were reported for this dosage in the other three trials.\textsuperscript{120-122}

**Etanercept**

One good\textsuperscript{124} and three fair\textsuperscript{125-129} placebo-controlled trials determined the efficacy and safety of etanercept for the treatment of plaque psoriasis. The good quality study was a 12 week RCT that enrolled 618 patients with plaque psoriasis.\textsuperscript{124} Patients were randomized to etanercept 50mg twice weekly or to placebo. At endpoint, 47 percent of patients on etanercept achieved a PASI 75 response, compared with 5 percent on placebo ($P < 0.0001$). Likewise, significantly more patients on etanercept than on placebo experienced improvements in fatigue (FACIT-F [functional assessment of chronic illness therapy fatigue] 5.0 vs. 1.9; $P < 0.0001$). Patients with an existing depression at baseline achieved a greater improvement of depressive symptoms when treated with etanercept than with placebo. The other three studies provided consistent results.\textsuperscript{125-129} In addition to better PASI scores, patients on etanercept achieved statistically significantly greater improvements of quality of life (DLQI, SF-36) than patients on etanercept.\textsuperscript{125-129}

**Infliximab**

Three studies assessed the efficacy and safety of infliximab for the treatment of plaque psoriasis.\textsuperscript{130,131,132} The only good RCT randomized 378 patients to infliximab (5mg/kg) or placebo.\textsuperscript{130} Patients were followed up for 24 weeks. In a 50 week extension study, those treated with placebo crossed over to infliximab treatment. At week 24, 82 percent of patients on infliximab and 4 percent of patients on placebo achieved a PASI 75 response ($P < 0.001$). In addition, the infliximab group had statistically significantly greater improvements on SF-36, DLQI\textsuperscript{133}, NAPSI (nail psoriasis and severity index), and PGA. The other two studies reported similar short term results at 10 weeks.\textsuperscript{131,132,134}

One trial reported that continuous maintenance regimens led to significantly better outcomes at 50 weeks than intermittent infliximab maintenance.\textsuperscript{132}
Table 12. Summary of Efficacy Trials in Adult Patients with Plaque Psoriasis

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Comparisons</th>
<th>Primary outcome</th>
<th>Secondary outcomes</th>
<th>Population</th>
<th>Results</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ALEFACEPT</strong></td>
<td>Ellis et al., 2001&lt;sup&gt;116, 118, 119&lt;/sup&gt;</td>
<td>RCT</td>
<td>229</td>
<td>24 weeks</td>
<td>ALE / placebo</td>
<td>PASI, Quality of life</td>
<td>Patients with plaque psoriasis without any systemic treatment</td>
<td>Significant improvement in PASI and quality of life outcomes for ALE</td>
<td>Fair</td>
</tr>
<tr>
<td></td>
<td>Lebwohl et al., 2003&lt;sup&gt;115, 117, 135&lt;/sup&gt;</td>
<td>RCT</td>
<td>507</td>
<td>24 weeks</td>
<td>ALE / placebo</td>
<td>PASI, PGA</td>
<td>Quality of life</td>
<td>Patients with plaque psoriasis without any systemic treatment</td>
<td>Significant improvement in PASI and quality of life outcomes for ALE</td>
</tr>
<tr>
<td><strong>ETANERCEPT</strong></td>
<td>Gottlieb et al., 2003&lt;sup&gt;127&lt;/sup&gt;</td>
<td>RCT</td>
<td>112</td>
<td>24 weeks</td>
<td>ETA / placebo</td>
<td>PASI</td>
<td>DLQI</td>
<td>Patients with plaque psoriasis without any systemic treatment</td>
<td>Significant improvement of PASI and DLQI for ETA</td>
</tr>
<tr>
<td></td>
<td>Papp et al., 2005&lt;sup&gt;128, 129&lt;/sup&gt;</td>
<td>RCT</td>
<td>583</td>
<td>12 weeks</td>
<td>ETA / placebo</td>
<td>DLQI, SF-36</td>
<td>PCS</td>
<td>Patients with plaque psoriasis without any systemic treatment</td>
<td>Significant improvement of DLQI and SFG-36 for ETA</td>
</tr>
<tr>
<td></td>
<td>Leonardi et al., 2003&lt;sup&gt;125, 126&lt;/sup&gt;</td>
<td>RCT</td>
<td>672</td>
<td>24 weeks</td>
<td>ETA / placebo</td>
<td>PASI 75</td>
<td>DLQI</td>
<td>Patients with plaque psoriasis without any systemic treatment</td>
<td>Significant improvement of PASI 75 and DLQI for ETA</td>
</tr>
<tr>
<td></td>
<td>Tyring et al., 2006&lt;sup&gt;124&lt;/sup&gt;</td>
<td>RCT</td>
<td>618</td>
<td>12 weeks</td>
<td>ETA / placebo</td>
<td>PASI, FACIT-F</td>
<td>Depression</td>
<td>Patients with plaque psoriasis without any systemic treatment</td>
<td>Significant improvement in PASI, fatigue, and depression for ETA</td>
</tr>
<tr>
<td><strong>EFALIZUMAB</strong></td>
<td>Gordon et al., 2003&lt;sup&gt;123, 136&lt;/sup&gt;</td>
<td>RCT</td>
<td>556</td>
<td>12 weeks</td>
<td>EFA / placebo</td>
<td>PASI 75</td>
<td>sPGA, DLQI, itching VAS, PSA</td>
<td>Patients with plaque psoriasis without any systemic treatment</td>
<td>Significantly greater improvement on all outcome measures for EFA than for placebo</td>
</tr>
<tr>
<td></td>
<td>Lebwohl et al., 2003&lt;sup&gt;120&lt;/sup&gt;</td>
<td>RCT</td>
<td>597</td>
<td>12 weeks</td>
<td>EFA / placebo</td>
<td>PASI</td>
<td>NR</td>
<td>Patients with plaque psoriasis without any</td>
<td>Significantly greater improvement on all</td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>N</td>
<td>Duration</td>
<td>Treatment</td>
<td>Outcome Measures</td>
<td>Patient Population</td>
<td>Quality Rating</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------</td>
<td>---</td>
<td>----------</td>
<td>-----------</td>
<td>------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leonardi et al., 2005&lt;sup&gt;121&lt;/sup&gt;</td>
<td>RCT</td>
<td>498</td>
<td>12 weeks</td>
<td>EFA / placebo</td>
<td>PASI 75, sPGA,</td>
<td>Patients with plaque psoriasis without any systemic treatment</td>
<td>Fair</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ortonne et al., 2005, CLEAR trial&lt;sup&gt;122&lt;/sup&gt;</td>
<td>RCT</td>
<td>526</td>
<td>12 weeks</td>
<td>EFA / placebo</td>
<td>SF-36, DLQI, PSA, itching VAS, PGPA</td>
<td>Patients with plaque psoriasis without any systemic treatment</td>
<td>Fair</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gottlieb et al. 2004&lt;sup&gt;131, 134&lt;/sup&gt;</td>
<td>RCT</td>
<td>249</td>
<td>10 weeks</td>
<td>INF / placebo</td>
<td>PASI, PGA, DLQI</td>
<td>Patients with plaque psoriasis without any systemic treatment</td>
<td>Fair</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Menter et al. 2006&lt;sup&gt;132&lt;/sup&gt;</td>
<td>RCT</td>
<td>835</td>
<td>50 weeks</td>
<td>INF/placebo</td>
<td>PASI, PGA, DLQI</td>
<td>Patients with plaque psoriasis without any systemic treatment</td>
<td>Fair</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reich et al., 2005&lt;sup&gt;130, 133&lt;/sup&gt;</td>
<td>RCT</td>
<td>378</td>
<td>24 weeks</td>
<td>INF / placebo</td>
<td>PASI, PGA, NAPSI, DLQI, SF-36</td>
<td>Patients with plaque psoriasis without any systemic treatment</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**ALE**: alefacept; **DLQI**: Dermatology Life Quality Index; **EFA**: efalizumab; **ETA**: etanercept; **INF**: infliximab; **NAPSI**: Nail Psoriasis and Severity Index; **PASI**: Psoriasis Area and Severity Index; **PGA**: Physician Global Assessment; **SF-36**: Medical Outcomes Study Short Form 36 Health Survey; **VAS**: Visual Analogue Scale
KEY QUESTION 2
What are the comparative incidence and severity of complications of included drugs?

A. Summary of the Evidence
The overall grade of the evidence on the comparative tolerability is poor. The only direct evidence on the comparative incidence of adverse events comes from one non-randomized, open-label trial comparing etanercept to infliximab in patients with RA. This 12-month study did not report any differences in tolerability. Evidence from placebo-controlled trials and observational studies is insufficient to draw conclusions about the comparative tolerability and safety of TIMs.

In efficacy studies TIMs were generally well tolerated. Injection site reactions (adalimumab, anakinra, efalizumab, etanercept) and infusion reactions (abatacept, infliximab, rituximab) were the most commonly and consistently reported adverse events. Some infusion reactions, however, appeared to be more serious than injection site reactions. One percent of patients had severe acute reactions that resembled acute anaphylactic conditions or led to convulsions. Injection site reactions were the most common reason for discontinuation due to adverse events. Incidence rates appear to be significantly higher with anakinra than with anti-TNF drugs.

Long-term, rare but serious adverse events such as malignancies, serious infections, or autoimmunity are a cause of concern for all TIMs and could not be assessed reliably in efficacy trials. Some observational studies indicate that infliximab might have a higher risk of granulomatous infections than etanercept. Hepatotoxicity has been reported for infliximab but not for other TIMs. The current evidence on rare but severe adverse events is limited to observational evidence such as case reports, database reviews, and open-label extension studies of RCTs which cannot reliably establish a causal relationship. Nevertheless, because of the absence of studies with the methodological strength to account for rare adverse events, even weak evidence may be important.

B. Overall Tolerability
Most studies that examined the general efficacy of TIMs also determined their tolerability. In addition, some RCTs had an open-label extension phase of up to three years. Methods of adverse events assessment, however, differed greatly. Few studies used objective scales such as the UKU-SES (Utvalg for Kliniske Undersogelser Side Effect Scale) or the adverse reaction terminology from the World Health Organization (WHO). Most studies combined patient-reported adverse events with a regular
clinical examination by an investigator. Often determining whether assessment methods were unbiased and adequate was difficult. Rarely were adverse events prespecified and defined. Short study durations and small sample sizes additionally limited the validity of adverse events assessment with respect to rare but serious adverse events.

Only three RCTs were designed to assess adverse events as primary outcomes. Most published studies assessing adverse events were post hoc analyses or retrospective reviews of databases. We included observational studies if the sample size was larger than 100 and the study duration was at least 1 year (Table 13).

Overall, TIMs appeared to have a good tolerability profile, although some rare but serious adverse events such as serious infections, lymphoma, leucopenia, or demyelinations are of concern. Discontinuation rates because of adverse events in patients treated with TIMs ranged from 3 to 16 percent and generally did not differ significantly from those in patients treated with placebo. A German retrospective, population-based cohort study reported that discontinuation rates because of adverse events, after 12 months of treatment were 16 percent for anakinra, 13 percent for etanercept, and 19 percent for infliximab.

Long-term extension studies of RCTs and safety analyses of postmarketing surveillance reported that the incidence of adverse events does not increase over time. A population-based post-marketing cohort study from Sweden reported that in 27 percent of patients treated with etanercept, at least one adverse event was reported. A two year open-label extension study of etanercept in children with JRA reports a serious adverse events rate of 16 percent, primarily due to infections.

Injection site reactions, abdominal pain, nausea, headache, diarrhea, upper respiratory tract infections, and urinary tract infections were the most commonly reported adverse events.

The only head-to-head study that we found for efficacy outcomes also assessed differences in tolerability and safety between etanercept and infliximab. This study used the adverse reaction terminology from the WHO to determine adverse events. Overall, no significant differences in adverse events were reported between etanercept and infliximab. The overall discontinuation rates at 20 months were also similar (etanercept 21%; infliximab 25%).
One large, multinational RCT was designed primarily to evaluate the safety of anakinra over 6 months. A total of 1,414 patients were randomized to anakinra (100 mg) or placebo. After 6 months the rate of adverse events did not differ significantly between anakinra and placebo, except for injection site reactions (72.6% vs. 32.9%; P-value not reported). Overall discontinuation rates (anakinra 21.6%; placebo 18.7%) and the rate of serious adverse events (anakinra 7.7%; placebo 7.8%) were also similar. However, a trend towards an increased risk of serious infections in anakinra-treated patients was apparent (2.1% vs. 0.4%; \( P = 0.068 \)). The STAR (Safety Trial of Adalimumab in Rheumatoid Arthritis) study determined the safety of adalimumab in combination with standard rheumatoid therapy. At 22 weeks, there were no significant differences between adalimumab and placebo with respect to adverse events.

Injection site reactions (adalimumab, anakinra, etanercept) and infusion reactions (infliximab) were the most commonly and consistently reported adverse events. Some infusion reactions, however, appeared to be more serious than injection site reactions. An observational study of 165 consecutive patients with Crohn’s disease reported that 8.4 percent of patients had infusion reactions to infliximab. These were mostly non-specific symptoms such as headache, dizziness, nausea, pruritus, chills, or fever. One percent of patients, however, had severe acute reactions that resembled acute anaphylactic conditions or led to convulsions. In clinical trials, 17 percent of patients experienced infusion reactions, 0.5 percent of those were severe. A prospective cohort study in patients with RA in a clinical care setting reported substantially higher numbers. In this study (n = 113 patients with 1183 infusions) 53 percent of patients experienced at least one infusion reaction during the course of the therapy (mean 15 months). Less than two percent of patients in clinical trials discontinued because of infusion reactions. In contrast, injection site reactions were mainly erythema, pruritus, rash, and pain of mild to moderate severity. However, injection site reactions were the most common reason for discontinuation due to adverse events. The mean, crude incidence of injection site reactions in RCTs and observational studies reviewed for this report was 17.5 percent (95% CI 7.1-27.9) for adalimumab, 22.4 percent (95% CI 8.5-36.3) for etanercept, but 67.2 percent (95% CI 38.7-95.7) for anakinra. The higher incidence of injection site reactions for anakinra over adalimumab and etanercept is consistent with numbers reported in the respective package inserts.

C. Specific Adverse Events

Serious Infections
Because of the immunosuppressive nature of TIMs, serious infections including tuberculosis, pneumonia, osteomyelitis, and sepsis are of special concern. The FDA has issued black box warnings about an
increased risk of infections for adalimumab and infliximab. The package inserts of anakinra and etanercept also contain warnings in bold letters.

In efficacy trials, the incidence of serious infections was consistently higher in TIM- than in placebo-treated patients. However, although clinically significant, differences rarely reached statistical significance due to lack of power. For example, in the large safety RCT (n = 1,414), a trend towards an increased risk of serious infections in anakinra-treated patients was apparent during the 6 months of treatment (2.1% vs. 0.4%; \( P = 0.068 \)). A good meta-analysis pooled data of more than 5,000 RA patients from adalimumab and infliximab efficacy trials. The pooled odds ratio for serious infections was 2.0 (95% CI 1.3 – 3.1). The number needed to harm (NNH) was 59 (95% CI 39 – 125) within a treatment period of 3 to 12 months.

The START (Trial for Rheumatoid Arthritis with Remicade) study was good RCT (n = 1,084) conducted to assess the risk of serious infections during infliximab treatment for RA. After 22 weeks of treatment patients on 3mg/kg infliximab had similar rates of serious infections as patients on placebo (RR: 1.0; 95% CI 0.3 – 3.1). Patients treated with 10mg/kg infliximab had a significantly higher rate of serious infections than patients on placebo (RR: 3.1 95% CI 1.2 – 7.9).

Most long-term observational studies support these findings. The most common serious infections were cases of tuberculosis. In addition, observational studies reported infections with coccidiomycosis, histoplasmosis, pneumocystis carinii, and listeriosis and candida.

Five retrospective database analyses and a prospective cohort study with a historic control group specifically determined the risk of tuberculosis or granulomatous infections during treatment with infliximab and etanercept. All studies report a significant increase of risk attributable to anti-TNF therapy. Three studies analyzed all reports of tuberculosis or granulomatous infections after infliximab or etanercept therapy through the MedWatch reporting system of the FDA. In general, the MedWatch system relies on voluntary reporting of adverse events and underreporting is likely. Therefore, it lacks an adequate denominator to draw inferences about causation and the comparative risks of any drugs. Among RA patients on infliximab, 24.4 cases of tuberculosis per 100,000 patients treated in the past year. In contrast, the estimated background rate for patients with RA not exposed to TIMs in the US is 6.2 cases per 100,000 patient years. Reported rates are lower than those of a prospective cohort study of patients from the National Data Bank for Rheumatic Diseases (NDP). This study reports 52.5 cases per 100,000 patients years. The median interval from start of infliximab therapy to the diagnosis of
tuberculosis was 3 months.\textsuperscript{139} By contrast, an analysis of MedWatch data concerning etanercept and tuberculosis reported a median time of 11.5 months from start of etanercept therapy to diagnosis of tuberculosis.\textsuperscript{162} The analysis of MedWatch data on granulomatous infections indicated a higher rate among patients treated with infliximab (239 cases per 100,000 patients) than with etanercept (74 cases per 100,000 patients).\textsuperscript{138} The rate of tuberculosis in this study was 144 cases per 100,000 patients for infliximab and 35 cases per 100,000 patients for etanercept. However, incidence rates are not comparable across studies because the Wallis et al. study reports cases per treated patients and not per patient years.\textsuperscript{138}

Two other database analyses used the Spanish BIOBADASER (Base de Datos de Productos Biologicos de la Sociedad Espanola de Reumatologia)\textsuperscript{167} and different Swedish databases\textsuperscript{163} which included data on infliximab and etanercept. Both reports indicated a substantially increased risk for tuberculosis in patients treated with etanercept or infliximab. The Swedish study reported a 4 – fold increased risk of tuberculosis (RR 4.0; 95% CI 1.3 – 12) for patients on anti-TNF treatment compared with RA patients not exposed to etanercept or infliximab.\textsuperscript{163}

In children on etanercept for JRA, the rate of serious infections over 4 years was 0.04 per patient-year.\textsuperscript{170}

**Lymphoma and other malignancies**

The risk of lymphoma, both Hodgkin and non-Hodgkin lymphoma, is generally increased in patients with RA.\textsuperscript{171} Data from controlled trials do not provide sufficient evidence concerning a further increase of risk attributable to TIMs or a combination of TIMs and MTX. A MedWatch report identified 26 reported cases of lymphoproliferative disorders in patients treated with infliximab or etanercept for Crohn’s disease or RA as of 2002.\textsuperscript{172} The estimated crude incidence rates of lymphoma are 19 per 100,000 patients treated with etanercept and 6.6 per 100,000 patients treated with infliximab. Authors report that in a number of cases, lymphoma developed shortly after starting therapy and regression occurred in two patients after discontinuing therapy. The median time from start of therapy until diagnosis was 8 weeks for etanercept and 6 weeks for infliximab. Given the fact that this study is essentially a case series, a clear causal relationship between TIMs and lymphoma, or differences in risk between drugs cannot be established.

A large prospective cohort study followed 18,572 RA patients registered in the National Data Bank of Rheumatic Diseases (NDB) for up to 3 years.\textsuperscript{173} Results indicated that lymphomas are increased in patients on anti-TNF-\(\alpha\) therapies. However, confidence intervals for treatment groups overlap and results are insufficient to establish a causal relationship between RA treatments and lymphoma or to delineate differences in risk between treatments. The standardized incidence rate (SIR) in the overall cohort was 1.9
cases per 100,000. The SIR for patients not receiving MTX or any biologic agents was 1.0. The SIR for patients on MTX was 1.7 (95%CI 0.9-3.2), on infliximab was 2.6 (95%CI 1.4-4.5), and on etanercept was 3.8 (95%CI 1.9-7.5). A Swedish retrospective cohort study of 1557 community-based patients also detected an increased, non-significant risk of lymphoma in patients on anti-TNF drugs compared with those on DMARDs (RR: 4.9; 95% CI 0.9 – 26.2).174

A good meta-analysis pooled data of more than 5,000 RA patients from adalimumab and infliximab efficacy trials.159 The pooled odds ratio for malignancies was 3.3 (95% CI 1.2 – 9.1). The number needed to harm (NNH) was 154 (95% CI 91 – 500) within a treatment period of 3 to 12 months. A large retrospective Swedish cohort study, however, does not support such a finding.175 This study, based on data of more than 60,000 patients, found similar standardized incidence ratios for solid cancers between RA patients treated with anti-TNF medications and those on conventional therapy using both a contemporary and a historic control.

A clinical trial database review did not detect an increased incidence of squamous cell carcinoma in 1,442 RA patients (4,257 patient years) treated with etanercept (crude rate: 2.8 cases/1000 patients).176 However, the median follow-up time was only 3.7 years. A larger retrospective cohort study (n = 15,789), however, reports a statistically significant association of a combination of anti-TNF treatment and MTX and non-melanoma skin cancer (HR: 1.28; 95% CI NR; P = 0.014).177

**Congestive Heart Failure**

No direct evidence on the comparative risk of TIMs for congestive heart failure (CHF) exists. The evidence on the risk of CHF with anti-TNF therapy is mixed. Two observational studies report lower rates of cardiovascular events178 and CHF179 for RA patients on anti-TNF therapy compared with those on conventional therapies. A good Swedish retrospective cohort study (n = 983), using data from population based databases, reported a statistically significantly lower risk of cardiovascular events in patients treated with anti-TNF medications compared with those on conventional therapy (age-sex adjusted rate ratio: 0.46/1000 person years; 95% CI 0.25 – 0.85; P = 0.013). A large retrospective cohort study (n = 13,171) based on the National Databank for Rheumatic Diseases reported an absolute risk reduction for CHF of 1.2 percent (95% CI -1.9 - -0.5; P = NR) for patients treated with anti-TNF therapy compared with those not treated with anti-TNF medications over a 2 year period.179 Confounding by indication, however, cannot entirely be ruled out with such study designs.
By contrast, a MedWatch analysis reports that half of the patients who developed new onset congestive heart failure (CHF) under etanercept or infliximab treatment did not have any identifiable risk factors.\footnote{180} Indirect evidences comes from three trials, two on etanercept\footnote{181} and one on infliximab,\footnote{182} that evaluated the efficacy of these drugs for the treatment of CHF. Information on the two etanercept studies, however, is limited to a review article.\footnote{181} The studies have not been published otherwise. We did not include this review article because it was not based on a systematic literature review. Nevertheless, we are briefly summarizing the findings.

Populations of these studies did not have any rheumatoid illnesses and, therefore, provide only indirect evidence. One of the two etanercept trials was terminated early because interim analyses indicated higher mortality rates in patients treated with etanercept. Similarly, the infliximab study presented higher mortality rates in the 10 mg/kg arm than in the placebo and 5 mg/kg arm.\footnote{182} The package insert of infliximab issues a contraindication regarding the use in patients with CHF; the package inserts of etanercept and adalimumab emphasize precaution.

**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, pancytopenia, and hepatotoxicity. Reports based on data from MedWatch indicated that adalimumab, etanercept, and infliximab might be associated with demyelination.\footnote{183,185} Similar cases have been seen in regulatory trials of adalimumab.\footnote{184} All neurologic events partially or completely resolved after discontinuation of treatment.

Similarly, reports of autoimmunity have not been confirmed in controlled trials and observational studies. However, case reports suggest an association between infliximab and drug induced lupus and other autoimmune diseases.\footnote{147,148,184,185} Lupus-like syndromes have also been reported for adalimumab.\footnote{150} A prospective cohort study of 125 consecutive Crohn’s disease patients treated infliximab reported a cumulative incidence of antinuclear antibodies of 56.8 percent after 24 months.\footnote{186} Two patients of this cohort developed drug induced lupus. Development of anti nuclear, anti double-stranded DNA, or anti-histone antibodies have also been reported in regulatory trials of other anti-TNF-α drugs.\footnote{156} The infliximab package insert reports that 34 percent of patients treated with infliximab and MTX experienced transient elevations of liver function parameters.\footnote{187} Severe liver injury, including acute liver failure has been reported. Owing to a lack of studies with the methodological strength to assess these rare
events, conclusions should be drawn on other grounds, such as comorbidities, taking case reports into consideration.

A prospective cohort study (n = 578) indicated that patients on anti-TNF treatments developed dermatological conditions (skin infections, eczema, drug-related eruptions) statistically significantly more often than anti-TNF naïve patients over a median treatment time of 2.3 years (25% vs. 13%, $P < 0.0005$).188
### Table 13. Summary of Studies Assessing Adverse Events

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Drug</th>
<th>Population</th>
<th>Results</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OVERALL TOLERABILITY</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Braun et al. 200579, 83-86</td>
<td>Open-label extension of RCT</td>
<td>70</td>
<td>3 years</td>
<td>INF</td>
<td>Patients with AS</td>
<td>INF is a well tolerated treatment</td>
<td>Fair</td>
</tr>
<tr>
<td>Cheifetz et al. 2003154</td>
<td>Case series</td>
<td>165</td>
<td>NR</td>
<td>INF</td>
<td>Patients with CD</td>
<td>Incidence of infusion reactions was 6.1%</td>
<td>N/A</td>
</tr>
<tr>
<td>Colombel et al. 2004147</td>
<td>Case series</td>
<td>500</td>
<td>median follow up 17 months</td>
<td>INF</td>
<td>Patients with CD</td>
<td>Incidence of serious adverse events was 8.6%</td>
<td>N/A</td>
</tr>
<tr>
<td>Feltelius et al. 2005153</td>
<td>Retrospective cohort study</td>
<td>1073</td>
<td>≥2 years</td>
<td>ETA</td>
<td>Patients with RA</td>
<td>27% of patients experienced at least one adverse event. The incidence of serious adverse events remained constant over time.</td>
<td>N/A</td>
</tr>
<tr>
<td>Fleischmann et al. 2003144-146</td>
<td>RCT</td>
<td>1,414</td>
<td>6 months</td>
<td>AKA</td>
<td>Patients with RA</td>
<td>AKA is a well tolerated treatment</td>
<td>Fair</td>
</tr>
<tr>
<td>Geneovese et al. 50</td>
<td>Open-label extension of RCT</td>
<td>201</td>
<td>5 years</td>
<td>ETA</td>
<td>Patients with RA</td>
<td>Higher rates of lymphoma compared to general population</td>
<td>Fair</td>
</tr>
<tr>
<td>Langer et al. 2003151</td>
<td>Case series</td>
<td>efficacy -166; safety 454</td>
<td>1 year</td>
<td>AKA</td>
<td>Patients with RA; population-based</td>
<td>Rate of adverse events was similar to those reported in efficacy trials</td>
<td>N/A</td>
</tr>
<tr>
<td>Ljung et al. 2005156</td>
<td>Case series</td>
<td>217</td>
<td>Up to 3 years</td>
<td>INF</td>
<td>Patients with IBD</td>
<td>19% experienced serious adverse events</td>
<td>N/A</td>
</tr>
<tr>
<td>Lovell et al. 200373, 142</td>
<td>Open-label extension of RCT</td>
<td>58</td>
<td>up to 2 years</td>
<td>ETA</td>
<td>Pediatric patients with polyarticular-JRA</td>
<td>16% of patients experienced serious adverse events</td>
<td>Fair</td>
</tr>
<tr>
<td>Lovell et al. 200670</td>
<td>Open-label extension of RCT</td>
<td>34</td>
<td>up to 4 years</td>
<td>ETA</td>
<td>Pediatric patients with polyarticular-JRA</td>
<td>Overall the rate of serious adverse events was 0.13 per patient-year</td>
<td>Fair</td>
</tr>
<tr>
<td>Maini et al. 200448, 69</td>
<td>Open-label extension of RCT</td>
<td>259</td>
<td>2 years</td>
<td>INF</td>
<td>Patients with RA</td>
<td>Rate of severe adverse events was similar in INF and placebo</td>
<td>Fair</td>
</tr>
<tr>
<td>Nuki et al. 2002152</td>
<td>Uncontrolled extension of RCT</td>
<td>309</td>
<td>76 weeks</td>
<td>AKA</td>
<td>Patients with RA</td>
<td>AKA was well tolerated at all dose levels for up to 76 weeks</td>
<td>N/A</td>
</tr>
<tr>
<td>Schaible et al. 2000148</td>
<td>Retrospective data analysis of clinical trials</td>
<td>913</td>
<td>12 weeks – 3 years</td>
<td>INF</td>
<td>Patients with CD or RA</td>
<td>Incidence of infections was greater in patients treated with INF than placebo</td>
<td>N/A</td>
</tr>
<tr>
<td>Study</td>
<td>Study Design</td>
<td>Sample Size</td>
<td>Follow-up</td>
<td>Treatment</td>
<td>Disease</td>
<td>Outcome Description</td>
<td>Methodological Quality</td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------------------------------</td>
<td>-------------</td>
<td>-----------</td>
<td>-----------</td>
<td>---------</td>
<td>---------------------------------------------------------------------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Schiff et al., 2006</td>
<td>Retrospective data analysis of clinical trials; postmarketing surveillance</td>
<td>10,050</td>
<td>12,506 patient years</td>
<td>ADA</td>
<td>Patients with RA</td>
<td>Long-term ADA treatment was generally safe</td>
<td>N/A</td>
</tr>
<tr>
<td>Wasserman et al., 2004</td>
<td>Prospective cohort study</td>
<td>113</td>
<td>15 months</td>
<td>INF</td>
<td>Patients with RA</td>
<td>53% of patients on INF experienced at least one infusion reaction.</td>
<td>Fair</td>
</tr>
<tr>
<td>Weinblatt et al., 2006</td>
<td>Open-label extension of RCT</td>
<td>162</td>
<td>3.4 years</td>
<td>ADA</td>
<td>Patients with RA</td>
<td>2.03 serious infections / 100 patient-years</td>
<td>Fair</td>
</tr>
<tr>
<td>Zink et al., 2005</td>
<td>Retrospective cohort study</td>
<td>1523</td>
<td>12 months</td>
<td>AKA, ETA, INF</td>
<td>Patients with RA</td>
<td>Similar discontinuation rates because of adverse events among AKA, ETA and INF</td>
<td>Fair</td>
</tr>
</tbody>
</table>

### INFECTION DISEASES

<table>
<thead>
<tr>
<th>Study</th>
<th>Study Design</th>
<th>Sample Size</th>
<th>Follow-up</th>
<th>Treatment</th>
<th>Disease</th>
<th>Outcome Description</th>
<th>Methodological Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salliot et al., 2006</td>
<td>Retrospective cohort study</td>
<td>709</td>
<td>NR</td>
<td>ADA, ETA, INF</td>
<td>Patients with rheumatic diseases; primary care-based cohort</td>
<td>The rates of serious infections in daily practice were higher than the ones reported in efficacy trials</td>
<td>N/A</td>
</tr>
<tr>
<td>Wolfe et al., 2006</td>
<td>Prospective cohort study</td>
<td>16,788</td>
<td>3.5 years</td>
<td>ADA, ETA, INF</td>
<td>Patients with RA</td>
<td>No increased risk for hospitalization for pneumonia foe ADA, ETA, and INF</td>
<td>Fair</td>
</tr>
<tr>
<td>Listing et al., 2005</td>
<td>Prospective cohort study</td>
<td>1529</td>
<td>Up to 12 months</td>
<td>AKA, ETA, INF</td>
<td>Patients with RA</td>
<td>Higher risk of infections for AKA, ETA, INF compared with DMARDs</td>
<td>Fair</td>
</tr>
<tr>
<td>Mohan et al., 2004</td>
<td>Database analysis Adverse Event Reporting System</td>
<td>25 cases</td>
<td>N/A</td>
<td>ETA</td>
<td>Patients treated with ETA</td>
<td>The median interval between first dose and diagnosis of tuberculosis was 11.5 months</td>
<td>N/A</td>
</tr>
<tr>
<td>Askling et al., 2005</td>
<td>Database analysis, Sweden</td>
<td>62,321</td>
<td>467,770 person years</td>
<td>ETA, INF</td>
<td>Patients with RA</td>
<td>4-fold increase of risk for tuberculosis for ETA and INF</td>
<td>N/A</td>
</tr>
<tr>
<td>Bergstrom et al., 2004</td>
<td>Retrospective cohort study</td>
<td>985</td>
<td>NR</td>
<td>ETA, INF</td>
<td>Patients with inflammatory arthritis</td>
<td>Patients treated with INF or ETA are more likely to develop symptomatic coccidioidomycosis</td>
<td>N/A</td>
</tr>
<tr>
<td>Bongartz et al., 2006</td>
<td>Meta-analysis</td>
<td>5014</td>
<td>3 to 12 months</td>
<td>ADA, INF</td>
<td>Patients with RA</td>
<td>Statistically significantly higher risk of serious infections for ADA and INF compared with placebo ($P = NR$)</td>
<td>Good</td>
</tr>
<tr>
<td>Gomez-Reino et al., 2003</td>
<td>Database analysis BIOBADASER</td>
<td>1540</td>
<td>Any duration</td>
<td>ETA, INF</td>
<td>Patients treated with INF or ETA</td>
<td>TB is more common in patients treated with INF or ETA</td>
<td>N/A</td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>Outcome measures</td>
<td>Patients treated</td>
<td>Adverse Events</td>
<td>Summary</td>
<td>Modality</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>----------------</td>
<td>---------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Lee et al. 2002</td>
<td>Database analysis</td>
<td>Histioplasmosis</td>
<td>ETA, INF</td>
<td>10 cases</td>
<td>Histioplasmosis infections may be a serious complication of treatment with anti-TNF agents; patients on INF had a higher rate of infections than patients on ETA</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Slifman et al. 2003</td>
<td>Database analysis</td>
<td>Listeria</td>
<td>ETA, INF</td>
<td>15 cases</td>
<td>Listeria infections may be a serious complication of treatment with anti-TNF agents; patients on INF had a higher rate of infections than patients on ETA</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Wallis et al. 2004</td>
<td>Database analysis</td>
<td>Patients on INF</td>
<td>ETA, INF</td>
<td>649 cases</td>
<td>Patients on INF had a higher rate of granulomatous infections than patients on ETA</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Keane et al. 2001</td>
<td>Database analysis</td>
<td>Patients treated</td>
<td>INF</td>
<td>70 cases</td>
<td>TB may develop soon after treatment with INF</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Lichtenstein et al., 2006</td>
<td>Prospective cohort</td>
<td>INF / Other</td>
<td>INF</td>
<td>Mean 1.9 years</td>
<td>Mortality rates and serious infections between INF and other therapies were similar</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Wolfe et al. 2004</td>
<td>Prospective cohort</td>
<td>Patients treated</td>
<td>INF</td>
<td>17,242</td>
<td>TB is more common in patients treated with INF</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Westhovens et al., 2006 (START)</td>
<td>RCT</td>
<td>INF + MTX / MTX</td>
<td>Patients treated</td>
<td>1084</td>
<td>The risk of serious infections was similar between placebo and 3mg/kg infliximab. 10mg/kg infliximab led to increased risk of serious infections.</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Askling et al., 2005</td>
<td>Retrospective cohort study</td>
<td>Anti-TNF</td>
<td>Patients with RA</td>
<td>60,930</td>
<td>No increase in solid cancers for patients treated with anti-TNF drugs</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Bongartz et al. 2006</td>
<td>Meta-analysis</td>
<td>ADA, INF</td>
<td>Patients with RA</td>
<td>5014</td>
<td>Statistically significantly higher risk of malignancies for ADA and INF compared with placebo</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Brown et al. 2002</td>
<td>Database analysis MedWatch</td>
<td>INF, ETA</td>
<td>Patients with RA or CD</td>
<td>26 cases</td>
<td>Estimated rate of lymphoma per 100,000 treated ETA- 19 INF- 6.6</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

**LYMPHOMA AND OTHER MALIGNANCIES**

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Outcome measures</th>
<th>Patients treated</th>
<th>Adverse Events</th>
<th>Summary</th>
<th>Modality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Askling et al., 2005</td>
<td>Retrospective cohort study</td>
<td>Anti-TNF</td>
<td>Patients with RA</td>
<td>60,930</td>
<td>No increase in solid cancers for patients treated with anti-TNF drugs</td>
<td>N/A</td>
</tr>
<tr>
<td>Bongartz et al. 2006</td>
<td>Meta-analysis</td>
<td>ADA, INF</td>
<td>Patients with RA</td>
<td>5014</td>
<td>Statistically significantly higher risk of malignancies for ADA and INF compared with placebo</td>
<td>Good</td>
</tr>
<tr>
<td>Brown et al. 2002</td>
<td>Database analysis MedWatch</td>
<td>INF, ETA</td>
<td>Patients with RA or CD</td>
<td>26 cases</td>
<td>Estimated rate of lymphoma per 100,000 treated ETA- 19 INF- 6.6</td>
<td>N/A</td>
</tr>
<tr>
<td>Study Authors and Year</td>
<td>Study Type</td>
<td>Participants</td>
<td>Follow-Up</td>
<td>Treatment(s)</td>
<td>Condition</td>
<td>Conclusion</td>
</tr>
<tr>
<td>------------------------</td>
<td>------------</td>
<td>--------------</td>
<td>-----------</td>
<td>--------------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>Chakravarty et al., 2005</td>
<td>Retrospective cohort study</td>
<td>15,789</td>
<td>NR</td>
<td>ETA, INF</td>
<td>Patients with RA</td>
<td>Statistically significant association between anti-TNF+MTX use and non-melanoma skin cancer</td>
</tr>
<tr>
<td>Flendrie et al., 2005</td>
<td>Prospective cohort study with historic control</td>
<td>578</td>
<td>911 patient years</td>
<td>Anit-TNF</td>
<td>Patients with RA</td>
<td>Higher rates of dermatological conditions in patients on anti-TNF drugs compared to DMARDs</td>
</tr>
<tr>
<td>Geborek et al., 2005</td>
<td>Retrospective cohort study</td>
<td>1557</td>
<td>5551 patient years</td>
<td>ETA, INF</td>
<td>Patients with RA</td>
<td>Higher risk of lymphoma for anti-TNF drugs</td>
</tr>
<tr>
<td>Lebwohl et al. 2005</td>
<td>Database review</td>
<td>1,442</td>
<td>3.7 years</td>
<td>ETA</td>
<td>Patients with RA</td>
<td>ETA does not seem to be associated with an increase in the incidence of cutaneous squamous cell carcinoma</td>
</tr>
<tr>
<td>Wolfe et al. 2004</td>
<td>Prospective cohort study</td>
<td>18,572</td>
<td>Up to 3 years</td>
<td>INF, ETA</td>
<td>Patients with RA</td>
<td>Patients with RA, treated with INF or ETA are more likely to develop lymphoma than the general population</td>
</tr>
</tbody>
</table>

**CONGESTIVE HEART FAILURE**

<table>
<thead>
<tr>
<th>Study Authors and Year</th>
<th>Study Type</th>
<th>Participants</th>
<th>Follow-Up</th>
<th>Treatment(s)</th>
<th>Condition</th>
<th>Conclusion</th>
<th>Quality Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chung et al. 2003</td>
<td>RCT</td>
<td>150</td>
<td>28 weeks</td>
<td>INF</td>
<td>Patients with CHF</td>
<td>INF (10mg ) -treated patients were more likely to die or have heart failure than placebo-treated patients</td>
<td>Fair</td>
</tr>
<tr>
<td>Jacobsson et al. 2005</td>
<td>Retrospective cohort study</td>
<td>983</td>
<td>NR</td>
<td>Anti-TNT Retrospective cohort study</td>
<td>Patients with RA</td>
<td>Patients on anti-TNF treatment had a lower rate of cardiovascular events than patients on traditional RA therapy</td>
<td>Good</td>
</tr>
<tr>
<td>Kwon et al. 2003</td>
<td>Database review MedWatch</td>
<td>47 cases</td>
<td>N/A</td>
<td>ETA, INF</td>
<td>Patients on ETA or INF therapy</td>
<td>Most patients with congestive heart failure did not have preexisting conditions</td>
<td>N/A</td>
</tr>
<tr>
<td>Wolfe et al. 2004</td>
<td>Retrospective cohort study</td>
<td>13,171</td>
<td>2 years</td>
<td>Retrospective cohort study</td>
<td>Patients with RA</td>
<td>Patients on anti-TNF treatment had a lower rate of congestive heart failure than patients on traditional RA therapy</td>
<td>Fair</td>
</tr>
</tbody>
</table>

**DEMYELINATION**

<table>
<thead>
<tr>
<th>Study Authors and Year</th>
<th>Study Type</th>
<th>Participants</th>
<th>Follow-Up</th>
<th>Treatment(s)</th>
<th>Condition</th>
<th>Conclusion</th>
<th>Quality Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mohan et al. 2001</td>
<td>Database analysis MedWatch</td>
<td>19 cases</td>
<td>N/A</td>
<td>Anti-TNF</td>
<td>Patients with inflammatory arthritis</td>
<td>All events temporally related to therapy, with partial or complete resolution on discontinuation.</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**AUTOIMMUNITY**

<table>
<thead>
<tr>
<th>Study Authors and Year</th>
<th>Study Type</th>
<th>Participants</th>
<th>Follow-Up</th>
<th>Treatment(s)</th>
<th>Condition</th>
<th>Conclusion</th>
<th>Quality Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vermeire et al 2003</td>
<td>Case series</td>
<td>125</td>
<td>Up to 24 months</td>
<td>INF</td>
<td>Patients with CD</td>
<td>AKA developed in 56.8% of treated patients</td>
<td>N/A</td>
</tr>
</tbody>
</table>
### OTHER ADVERSE EVENTS

<table>
<thead>
<tr>
<th>Study</th>
<th>Study Type</th>
<th>Patients</th>
<th>Median Follow-Up</th>
<th>Drug</th>
<th>Patients with Condition</th>
<th>Description</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baeten et al. 2003</td>
<td>Case series</td>
<td>107</td>
<td>191.5 patient years</td>
<td>INF</td>
<td>Patients with spondyloarthropathy</td>
<td>Though use of INF is generally safe care must be taken for serious adverse events such as infections and TB.</td>
<td>N/A</td>
</tr>
<tr>
<td>Colombel et al. 2004</td>
<td>Case series</td>
<td>500</td>
<td>median follow up 17 months</td>
<td>INF</td>
<td>Patients with CD</td>
<td>Short- and long-term INF therapy is generally well tolerated</td>
<td>N/A</td>
</tr>
</tbody>
</table>

ADA: adalimumab; AKA: anakinra; AS: ankylosing spondylitis; CD: Crohn’s disease; DMARD: disease-modifying antirheumatic drug; ETA: etanercept; IBS: irritable bowel disease; INF: infliximab; JRA: juvenile rheumatoid arthritis; MTX: methotrexate; N/A: not applicable; RA: rheumatoid arthritis; RCT: randomized controlled trial; TB: tuberculosis; TNF: tumor necrosis factor
**KEY QUESTION 3**

Do the included drugs differ in effectiveness or adverse events in different age, sex, or ethnic groups, or in patients taking other commonly prescribed drugs?

**A. Summary of the Evidence**

The overall grade of the evidence on efficacy and tolerability in subgroups is poor. We did not identify any study specifically designed to compare the effect of adalimumab, alefacept, anakinra, efalizumab, etanercept, or infliximab in one subgroup of patients compared to another. Subgroup analyses and indirect evidence from placebo-controlled trials provide evidence for some TIM drugs.

Indirect evidence exists from five retrospective analyses\(^{191-195}\) that age is not associated with greater clinical response rates in AS, RA PsA, or plaque psoriasis. No differences in adverse events between patients older than 65 years and those younger were reported.\(^{193-195}\) In one prospective cohort study significantly more females than males developed antinuclear antibodies when treated with infliximab.\(^{186}\) A pooled analysis of nine efficacy studies of alefacept did not detect any differences in efficacy and safety for obese or diabetic patients with plaque psoriasis.\(^{195}\)

No direct evidence on the comparative risk of TIMs in patients with a condition of interest and congestive heart failure (CHF) exists. The evidence on the general risk of CHF with anti-TNF therapy is mixed. Observational studies indicate a lower risk\(^{178, 179}\), indirect evidence from RCTs suggests a higher risk for CHF. Indirect evidence from three RCTs(two publications)\(^{181, 182}\) conducted in patients with CHF indicates that treatment with etanercept and infliximab significantly increases the risk of hospitalization and mortality.

**B. Age**

We did not identify any study specifically designed to compare the effect of adalimumab, alefacept, anakinra, efalizumab, etanercept, or infliximab in a younger versus an older population.

The best evidence for plaque psoriasis stems from a pooled analysis of nine efficacy studies of alefacept for the treatment of plaque psoriasis. Patients older than 65 years did not show any differences in efficacy and safety compared to younger patients during 12 weeks of treatment.\(^{195}\)
Other evidence on the role of age as an effect modifier is mixed. Indirect evidence exists from four retrospective analyses\textsuperscript{191-194} that age is not associated with greater clinical response rates in AS, RA, and PsA. A large observational study indicated that response to treatment in RA patients treated with etanercept and infliximab was better in those younger than 65 years.\textsuperscript{25} A case series in patients with Crohn’s Disease reports that young age was associated with an increased short-term response.\textsuperscript{191} No differences in adverse events between patients with AS, RA, PsA, and plaque psoriasis older than 65 years and those younger were reported.\textsuperscript{193-195} However, selection bias might have distorted results in some of these retrospective analyses.

**C. Ethnicity**

We did not identify any study specifically designed to compare the effect of adalimumab, alefacept, anakinra, efalizumab, etanercept, or infliximab in one racial group compared to another. In general, trials were conducted predominantly in white populations. No indirect evidence suggests that effectiveness or adverse events differ among races.

**D. Sex**

We did not identify any study specifically designed to compare the effect of adalimumab, alefacept, anakinra, efalizumab, etanercept, or infliximab in females compared to males. On average, study populations comprised more females than males; this fact reflects population and disease demographics and does not provide insight into treatment differences. One prospective cohort study reported that significantly more women than men developed antinuclear antibodies under infliximab (OR 2.5; 95%CI 1.2-5.4).\textsuperscript{186} No other indirect evidence suggests that effectiveness or adverse events differ between females and males.

**E. Comorbidities**

We did not identify any study specifically designed to assess the efficacy of adalimumab, alefacept, anakinra, efalizumab, etanercept, or infliximab in patients with comorbidities.

Results from a pooled analysis of nine efficacy studies of alefacept for the treatment of plaque psoriasis indicated that alefacept has similar efficacy and safety in obese and diabetic patients compared to patients without these comorbidities.\textsuperscript{195}
A post hoc subgroup analysis of a large safety trial determined the safety profile of anakinra in patients with comorbidities (cardiovascular events, pulmonary events, diabetes, infections, malignancies, renal impairment, central nervous system-related events). \(^\text{146}\) Overall, the incidence rates of adverse events were similar regardless of comorbidity status.

No direct evidence on the comparative risk of TIMs in patients with an indication of interest and congestive heart failure (CHF) exists. The evidence on the risk of CHF with anti-TNF therapy is mixed. Two observational studies report lower rates of cardiovascular events \(^\text{178}\) and CHF \(^\text{179}\) for RA patients on anti-TNF therapy compared with those on conventional therapies. A good Swedish retrospective cohort study (n = 983), using data from population based databases, reported a statistically significantly lower risk of cardiovascular events in patients treated with anti-TNF medications compared with those on conventional therapy (age-sex adjusted rate ratio: 0.46/1000 person years; 95% CI 0.25 – 0.85; \(P = 0.013\)). A large retrospective cohort study (n = 13,171) based on the National Databank for Rheumatic Diseases reported an absolute risk reduction for CHF of 1.2 percent (95% CI -1.9 - -0.5; \(P = \text{NR}\)) for patients treated with anti-TNF therapy compared with those not treated with anti-TNF medications over a 2 year period. \(^\text{179}\) Confounding by indication, however, cannot entirely be ruled out with such study designs.

By contrast, indirect evidence exists regarding an increased risk of worsening heart failure and mortality during anti-TNF-α therapy. Three trials, two on etanercept \(^\text{181}\) and one on infliximab \(^\text{182}\) evaluated the efficacy of these drugs for the treatment of CHF. None of the patients had any rheumatoid illnesses. The two etanercept trials were terminated early because interim analyses indicated higher mortality rates in patients treated with etanercept. Similarly, the infliximab study presented higher mortality rates in the 10 mg/kg arm than in the placebo and 5 mg/kg arm. \(^\text{182}\) A MedWatch analysis reported that half of the patients who developed new onset CHF while treated with etanercept or infliximab for RA or other rheumatoid illnesses did not have any identifiable risk factors. \(^\text{180}\) The package insert of infliximab issues a contraindication regarding its use in patients with CHF; the package inserts of etanercept and adalimumab express precaution.

F. Other Subgroups

A case series of 131 pregnant women exposed to infliximab did not detect adverse pregnancy outcomes compared to the general population. \(^\text{196}\) However, the sample size of this study was small and limitations of case series must be kept in mind. In addition, 27 percent of patients were lost to follow-up.
G. Other Commonly Prescribed Medications

No formal drug interaction studies have been performed with adalimumab, alefacept, anakinra, efalizumab, etanercept, or infliximab. Concurrent administration of anakinra with TNF-blocking agents (i.e., adalimumab, etanercept, infliximab) may be associated with an increased risk of serious infections, an increased risk of neutropenia, and no additional benefit compared to monotherapy. This evidence comes from a 24 week trial comparing concurrent treatment with anakinra and etanercept to etanercept monotherapy in patients with RA.\textsuperscript{32} Patients treated with both anakinra and etanercept had a 7 percent rate of serious infections, compared to no infections observed in patients treated with etanercept alone. Two percent of patients treated concurrently with anakinra and etanercept developed neutropenia. Because adalimumab and infliximab have a similar mechanism of action to etanercept, similar risks are believed to be associated with concurrent treatment with anakinra, although no formal evidence exists.

Because the majority of patients included in clinical studies received one or more concomitant medications (e.g., ASAs, antibiotics, antivirals, azathioprine, corticosteroids, folic acid, narcotics, nonsteroidal anti-inflammatory agents, and 6-MP) with no identifiable differences in safety or tolerability, concomitant treatment with such agents is believed to be safe. One analysis of data from the first 6 months of a large, blinded, placebo-controlled safety trial of anakinra provides evidence for the risk of infections or other serious adverse events for some concomitant medications.\textsuperscript{145} In this trial, no statistically significant differences were noted in the risk of infection or other serious adverse events between placebo- and anakinra-treated patients concurrently taking MTX or other DMARDs. Two patients taking anakinra and azathioprine developed serious infections compared to no patients taking azathioprine and placebo, although the number of patients taking azathioprine was deemed to be too small to draw any definitive conclusions. The adverse event profiles were similar for anakinra and placebo for patients who were or were not taking concomitant antihypertensive, antidiabetic, or statin drugs.

Concomitant administration of adalimumab and MTX has demonstrated a 29 to 44 percent reduction in the clearance of adalimumab. However, data do not suggest the need for dose adjustment of either MTX or adalimumab. Studies evaluating concomitant administration of MTX with anakinra or etanercept have not demonstrated changes in the clearance either drug. Although no formal studies have evaluated drug
interactions between MTX and alefacept, efalizumab, or infliximab, concomitant administration of these agents is believed to be safe.
# Table 14. Summary of Studies Assessing Subgroups

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Drug</th>
<th>Population</th>
<th>Results</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AGE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fleischmann et al. 2005(^{193})</td>
<td>Retrospective data analysis</td>
<td>4322</td>
<td>NR</td>
<td>ETA</td>
<td>Patients with RA, AS, PsA</td>
<td>No differences in adverse events between patients older and younger than 65 years</td>
<td>N/A</td>
</tr>
<tr>
<td>Fleischmann et al. 2003(^{194})</td>
<td>Retrospective data analysis</td>
<td>1128</td>
<td>NR</td>
<td>ETA</td>
<td>Patients with RA</td>
<td>No differences in efficacy and adverse events between patients older and younger than 65 years</td>
<td>N/A</td>
</tr>
<tr>
<td>Gottlieb et al. 2005(^{195})</td>
<td>Pooled analysis of efficacy trials</td>
<td>NR</td>
<td>12 weeks</td>
<td>ALE</td>
<td>Patients with plaque psoriasis</td>
<td>No differences in efficacy and adverse events between patients older and younger than 65 years</td>
<td>Fair</td>
</tr>
<tr>
<td>Rudwaleit et al. 2004(^{192})</td>
<td>Retrospective data analysis</td>
<td>99</td>
<td>12 weeks</td>
<td>ETA, INF</td>
<td>Patients with AS</td>
<td>Age not statistically significantly associated with treatment response</td>
<td>N/A</td>
</tr>
<tr>
<td>Vermeire et al. 2002(^{191})</td>
<td>Case series</td>
<td>240</td>
<td>4-10 weeks</td>
<td>INF</td>
<td>Patients with CD</td>
<td>Young age favored short term response to INF therapy</td>
<td>N/A</td>
</tr>
<tr>
<td>Weaver et al. 2006(^{22})</td>
<td>Prospective cohort study</td>
<td>3694</td>
<td>52 weeks</td>
<td>ETA, INF</td>
<td>Patients with RA</td>
<td>Patients younger than 65 years had better response</td>
<td>Fair</td>
</tr>
<tr>
<td><strong>COMORBIDITIES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chung et al. 2003(^{182})</td>
<td>RCT</td>
<td>150</td>
<td>28 weeks</td>
<td>INF</td>
<td>Patients with CHF</td>
<td>INF-treated (10mg) patients were more likely to die or have heart failure than placebo-treated patients</td>
<td>Fair</td>
</tr>
<tr>
<td>Jacobsson et al. 2003(^{178})</td>
<td>Retrospective cohort study</td>
<td>983</td>
<td>NR</td>
<td>Anti-TNF</td>
<td>Patients with RA</td>
<td>Patients on anti-TNF treatment had a lower rate of cardiovascular events than patients on traditional RA therapy</td>
<td>Good</td>
</tr>
<tr>
<td>Gottlieb et al. 2005/195</td>
<td>Pooled analysis of efficacy trials</td>
<td>NR</td>
<td>12 weeks</td>
<td>ALE</td>
<td>Patients with plaque psoriasis</td>
<td>No differences in efficacy and adverse events in diabetic and obese patients compared to the general study population</td>
<td>Fair</td>
</tr>
<tr>
<td>Kwon et al. 2003(^{180})</td>
<td>Database review MedWatch</td>
<td>47</td>
<td>N/A</td>
<td>ETA, INF</td>
<td>Patients on ETA or INF therapy</td>
<td>Young age was associated with a greater short term response</td>
<td>N/A</td>
</tr>
<tr>
<td>Schiff et al. 2004(^{144, 146})</td>
<td>Subgroup analyses of RCT</td>
<td>1,414</td>
<td>6 months</td>
<td>AKA</td>
<td>Patients with RA</td>
<td>Incidence rates of adverse events similar in patients with comorbidities</td>
<td>Fair</td>
</tr>
<tr>
<td>Wolfe et al. 2004(^{179})</td>
<td>Retrospective cohort study</td>
<td>13,171</td>
<td>2 years</td>
<td>Retrospective cohort</td>
<td>Patients with RA</td>
<td>Patients on anti-TNF treatment had a lower rate of congestive heart failure than patients on traditional RA therapy</td>
<td>Fair</td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>N</td>
<td>Duration</td>
<td>Medication</td>
<td>Subgroup</td>
<td>Adverse Events</td>
<td>GRADE</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>--------</td>
<td>----</td>
<td>----------</td>
<td>------------</td>
<td>-------------------</td>
<td>----------------------------------------------------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Genovese et al., 2004</td>
<td>RCT</td>
<td>242</td>
<td>24 weeks</td>
<td>AKA + ETA</td>
<td>Patients with RA</td>
<td>Patients treated with both AKA and ETA had a 7 percent rate of serious infection, compared to no infections observed with ETA alone.</td>
<td>Fair</td>
</tr>
<tr>
<td>Tesser et al., 2004</td>
<td>RCT</td>
<td>1399</td>
<td>6 months</td>
<td>AKA</td>
<td>Patients with RA</td>
<td>The adverse event profiles were similar for anakinra and placebo for patients who were or were not taking concomitant antihypertensives, antidiabetic, or statin drugs.</td>
<td>Fair</td>
</tr>
<tr>
<td>Vermeire et al, 2003</td>
<td>Case</td>
<td>125</td>
<td>24 months</td>
<td>INF</td>
<td>Patients with CD</td>
<td>More women than men developed anti-nuclear antibodies under INF.</td>
<td>N/A</td>
</tr>
<tr>
<td>Katz et al., 2004</td>
<td>Case</td>
<td>131</td>
<td>NR</td>
<td>INF</td>
<td>Pregnant women on INF</td>
<td>Pregnancy results in women on INF did not differ from the general population.</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Notes:**
- AKA: anakinra; ALE: alefacept; AS: ankylosing spondylitis; CD: Crohn’s disease; CHF: congestive heart failure; ETA: etanercept; INF: infliximab; MTX: methotrexate; N/A: not applicable; NR: not reported; PsA: psoriatic arthritis; RA: rheumatoid arthritis; RCT: randomized controlled trial; TNF: tumor necrosis factor.

**Concomitant Medications**

**SEX**

**OTHER SUBGROUPS**
CONCLUSIONS

Insufficient evidence exists to draw firm conclusions about the comparative efficacy, effectiveness, or tolerability of abatacept, adalimumab, alefacept, anakinra, efalizumab, etanercept, infliximab, and rituximab for the treatment of RA, JRA, AS, PsA, Crohn’s disease, UC, and plaque psoriasis. No double-blind randomized trial compared one TIM to another. The only direct comparative evidence comes from one open-label effectiveness trial comparing the effectiveness of etanercept to infliximab for the treatment of RA. Although this trial did not detect any differences in effectiveness after one year, the study design cannot completely rule out bias and confounding. Adjusted indirect comparisons suggest that anakinra is less efficacious than anti-TNF drugs for the treatment of RA.

The general efficacy of abatacept, adalimumab, anakinra, etanercept, infliximab, and rituximab for the treatment of RA is well established by multiple good to fair RCTs. Effect sizes are large and consistent across studies. Combination therapy with MTX achieved the best results. Monotherapies of adalimumab and etanercept did not reveal a relative benefit to MTX monotherapy. However, radiographic outcomes were statistically significantly better in patients on TIMs than on MTX. It remains unclear however, whether such differences are clinically relevant and lead a different progression of the disease. Other TIMs have not been directly compared to MTX. A combination of two TIMs (i.e., etanercept and anakinra) did not raise response or remission rates but significantly increased adverse events.

Likewise, fair to good evidence exists about the general efficacy of alefacept, efalizumab, etanercept, and infliximab for the treatment of plaque psoriasis. Although effect sizes with respect to PASI response rates vary substantially among these drugs, no indirect comparisons should be made across individual placebo-controlled trials.

Evidence on the general efficacy of TIMs for other reviewed indications is limited. Fair evidence exists that etanercept and infliximab are more efficacious than placebo for the treatment of AS and PsA. Multiple good to fair RCTs confirm the efficacy of infliximab for the treatment of Crohn’s disease. Etanercept did not significantly improve symptoms of Crohn’s disease compared to placebo; however, this finding is limited to one study. Fair evidence exists from three studies that infliximab is more efficacious than placebo in the treatment of ulcerative colitis. JRA is the indication with the sparsest evidence on the efficacy and tolerability of TIMs. Only one RCT provides evidence on the efficacy of
etanercept, the only drug approved for the treatment of JRA; however, methodological issues limit the internal validity of this study. Results of an uncontrolled trial of infliximab for JRA are fatally flawed.

Overall, no substantial differences in short-term tolerability and safety appear to exist among TIMs. The existing evidence suggests that differences in short-term tolerability exist primarily with respect to adverse events caused by the route of administration. Anakinra appears to have 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.

Rare but severe adverse events such as serious infections, lymphoma, autoimmunity, or congestive heart failure are of concern for all drugs. Existing evidence is insufficient to draw firm conclusions about the comparative safety among TIMs. Because TIMs are relatively new medications, solid long-term data on safety is generally still missing.

The most obvious differences that might be clinically decisive for choosing a TIM involve dosing and administration. Abatacept, infliximab, and rituximab require intravenous administration at different intervals and present the danger of rare but severe infusion reactions. Adalimumab, anakinra, efalizumab, and etanercept can be administered subcutaneously by the patient. Administration intervals, however, differ substantially: adalimumab requires an injection once a week or once every other week, anakinra has to be administered daily, and etanercept and efalizumab once or twice per week. Alefacept requires an intramuscular injection.

Overall, TIMs are highly effective medications for the treatment of RA, JRA, AS, PsA, Crohn’s disease, UC, and plaque psoriasis that substantially improve the burden of disease. However, the risk benefit ratio cannot be reliably assessed without sound long-term data on safety.

**Gaps in the Evidence**
No well-conducted double-blind randomized head-to-head trials exist comparing one TIM with another. Evidence from systematic reviews, placebo-controlled trials, and observational studies is insufficient to draw firm conclusions about one TIM compared to another.
In addition, the lack of sound evidence for the treatment of JRA with TIMs is apparent. Currently, published studies do not have the methodological rigor required to assess the risk benefit ratio of TIM-therapy in a pediatric population.

Given the danger of severe, potentially fatal adverse events, large, long-term, well-conducted, observational studies are paramount to reliably assessing the risk benefit ratio of TIM-therapy. Future research should focus on prospectively evaluating the risk of rare but severe adverse events employing adequate study designs.
### Table 15. Summary of the Evidence

<table>
<thead>
<tr>
<th>Key Question 1: Comparative Efficacy</th>
<th>Rating of the Body of Evidence</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA</td>
<td>Fair-Poor</td>
<td>Only one non-randomized, open-label trial and four observational studies provide direct evidence on the comparative efficacy of etanercept and infliximab; etanercept had numerically greater ACR response rates than infliximab. Indirect comparisons of placebo controlled trials did not find statistically significant differences in efficacy among individual drugs. However, point estimates favor adalimumab, etanercept, and infliximab over anakinra. Adjusted indirect comparisons of anakinra with anti-TNF drugs as a class present a statistically significantly greater efficacy for anti-TNF drugs on ACR 20 but not on ACR 50. Multiple placebo-controlled trials provide good to fair evidence on the general efficacy of abatacept, adalimumab, anakinra, etanercept, infliximab, or rituximab for the treatment of RA.</td>
</tr>
<tr>
<td>JRA</td>
<td>Poor</td>
<td>We identified no head-to-head trials. The evidence for JRA is limited to one fair placebo-controlled trial establishing the efficacy of etanercept for the treatment of JRA.</td>
</tr>
<tr>
<td>AS</td>
<td>Poor</td>
<td>We identified no head-to-head trials. Five placebo-controlled trials provide good to fair evidence on the general efficacy of etanercept and infliximab for the treatment of AS. Significant differences in study characteristics make this evidence insufficient to identify differences among treatments. No studies on adalimumab, alefacept, anakinra, or efalizumab could be detected.</td>
</tr>
<tr>
<td>PsA</td>
<td>Poor</td>
<td>We identified no head-to-head trials. Six placebo-controlled trials provide fair evidence on the general efficacy of adalimumab, alefacept, etanercept and infliximab for the treatment of PsA. Significant differences in study characteristics make this evidence insufficient to identify differences among treatments. No studies on adalimumab, abetacept, anakinra, efalizumab, or rituximab could be detected.</td>
</tr>
<tr>
<td>Crohn’s Disease</td>
<td>Poor</td>
<td>We identified no head-to-head trials. Six placebo-controlled trials provide fair evidence on the general efficacy of infliximab for the treatment of Crohn’s disease. One short but good quality trial found adalimumab to be more efficacious than placebo. One fair trial could not detect any significant differences in efficacy between etanercept and placebo. Data was insufficient to conduct statistical indirect comparisons. No studies on abatacept, alefacept, anakinra, efalizumab, or rituximab could be detected.</td>
</tr>
<tr>
<td>Ulcerative Colitis</td>
<td>Poor</td>
<td>We identified no head-to-head trials. Three placebo-controlled trials provide fair evidence on the general efficacy of infliximab for the treatment of ulcerative colitis. No studies on abatacept, adalimumab, alefacept, anakinra, efalizumab, etanercept or rituximab could be detected.</td>
</tr>
</tbody>
</table>
Table 15: Summary of the Evidence

| Plaque Psoriasis | Poor | We identified no head-to-head trials. Twelve placebo-controlled trials provide fair evidence on the general efficacy of alefacept, efalizumab, etanercept, and infliximab for the treatment of plaque psoriasis. Significant differences in study characteristics make this evidence insufficient to identify differences among treatments. No studies on abatacept, adalimumab, anakinra, or rituximab could be detected. |

<table>
<thead>
<tr>
<th>Key Question 2: Comparative Adverse Events</th>
<th>Rating of the Body of Evidence</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tolerability and discontinuation</td>
<td>Fair to Poor</td>
<td>Only one non-randomized, open-label trial provides direct evidence on the comparative tolerability of etanercept and infliximab; no differences were apparent. Overall, the incidence rates of adverse events appear to be similar among reviewed TIMs. Anakinra appears to have a higher rate of injection site reactions than adalimumab and etanercept. Abatacept, infliximab and rituximab can cause severe infusion reactions. Infliximab has a potential for hepatotoxicity that has not been reported for other TIMs. Discontinuation rates because of adverse events did not differ significantly compared to placebo, taking the whole body of evidence into consideration.</td>
</tr>
<tr>
<td>Serious infections</td>
<td>Poor</td>
<td>Fair evidence from controlled trials and observational studies suggests that the rate of serious infections is higher for TIMs than for placebo. In particular, a higher risk of tuberculosis is well documented. Observational studies report increased infections with histioplasmosis, pneumocystis carinii, listeriosis or candida. Evidence from controlled trials and observational studies is insufficient to draw conclusions about the comparative risk of serious infections.</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>Poor</td>
<td>Observational evidence indicates a higher risk of lymphoma for patients treated with infliximab or etanercept. Evidence from controlled trials and observational studies is insufficient to draw conclusions about the comparative risk of lymphoma.</td>
</tr>
<tr>
<td>CHF</td>
<td>Poor</td>
<td>The evidence on the risk for CHF is mixed. Three RCTs provide fair, indirect evidence about a higher rate of mortality for patients with CHF treated with etanercept or infliximab than with placebo. Evidence from observational studies in patients with RA reported a lower risk of CHF for patients on anti-TNF drugs.</td>
</tr>
<tr>
<td>Demyelination</td>
<td>Poor</td>
<td>Case reports indicate that etanercept and infliximab might be associated with demyelination. Evidence, however, is insufficient to draw conclusions about differences in the risk of demyelination.</td>
</tr>
<tr>
<td>Autoimmunity</td>
<td>Poor</td>
<td>Case reports indicate that TIMs might be associated drug induced lupus and other forms of autoimmunity. Evidence from controlled trials and observational studies is insufficient to draw conclusions about differences in the risk of autoimmunity.</td>
</tr>
<tr>
<td>Neutropenia</td>
<td>Poor</td>
<td>One trial indicates that a combination of anakinra and etanercept is associated with an increased risk of panzytopenia. Evidence from controlled trials and observational studies is insufficient to draw conclusions about differences in the risk for panzytopenia</td>
</tr>
</tbody>
</table>
Table 15: Summary of the Evidence

<table>
<thead>
<tr>
<th>Hepatotoxicity</th>
<th>Poor</th>
<th>Evidence from controlled trials and observational studies is insufficient to draw conclusions about differences in the risk of liver toxicity.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key Question 3: Subgroups</td>
<td>Rating of the Body of Evidence</td>
<td>Conclusion</td>
</tr>
<tr>
<td>Age</td>
<td>Poor</td>
<td>A pooled analysis did not find a difference in the efficacy of alefacept in patients older and younger than 65 years treated for plaque psoriasis. Indirect evidence suggests that young age is associated with increased clinical response rates for patients with Crohn’s disease or AS. Evidence is insufficient to draw conclusions about age and differences in treatment effects among TIMs.</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Poor</td>
<td>Evidence is insufficient to draw conclusions about ethnicity and differences in treatment effects among TIMs.</td>
</tr>
<tr>
<td>Sex</td>
<td>Poor</td>
<td>Evidence is insufficient to draw conclusions about sex and differences in treatment effects among TIMs.</td>
</tr>
<tr>
<td>Comorbidities</td>
<td>Poor</td>
<td>Alefacept had a similar efficacy in diabetic and obese patients compared to patients without these conditions. We could not find any studies comparing the efficacy and tolerability of other TIMs between a population with a comorbidity and one without the same comorbidity. The evidence is mixed on a higher mortality in patients with CHF. Evidence is insufficient to draw conclusions about comorbidities and differences in treatment effects among TIMs.</td>
</tr>
</tbody>
</table>

CHF: congestive heart failure; RCT: randomized controlled trial; TIMs: targeted immune modulators; TNF: tumor necrosis factor
Figure 2. Results of literature search

- Titles and abstracts identified through searches: n = 2,849
- Articles published as abstract-only: n = 41
- Full-text articles retrieved: n = 504
- Full text articles excluded: n = 173
  - 20 Wrong outcomes
  - 22 Drug not included
  - 7 Population not included
  - 34 Wrong publication type
  - 90 Wrong study design
- Background articles: n = 165
- Articles included in drug class review: n = 166
  - 1 head-to-head non-randomized trial
  - 88 on placebo controlled trials
  - 5 on systematic reviews or meta-analyses
  - 51 on observational studies
  - 4 on pooled data analyses
  - 2 studies of included meta-analyses that were abstracted for adverse events
  - 11 on studies included in abstracted meta-analyses
  - 1 on study included in the meta-analysis only
  - 3 on studies deemed to be of poor quality

*Number of included articles differs from number of included studies due to the fact that some studies have multiple publications.
Appendix A. Search Strategy

#1 Search "infliximab"[Substance Name] OR remicade OR "TNFR-Fc fusion protein"[Substance Name] OR etanercept OR enbrel OR "adalimumab"[Substance Name] OR humira

#2 Search "interleukin 1 receptor antagonist protein"[Substance Name] OR kineret OR anakinra OR "efalizumab"[Substance Name] OR raptiva OR "alefacept"[Substance Name] OR amevive

#3 Search "cytotoxic T lymphocyte-associated antigen 4-immunoglobulin"[Substance Name] OR Abatacept OR orencia OR "rituximab"[Substance Name] OR rituxan OR anti-cd20a

#4 Search #1 OR #2 OR #3

#5 Search ("Arthritis, Rheumatoid"[MeSH] OR "Arthritis, Juvenile Rheumatoid"[MeSH]) OR ankylosing arthritis OR "Arthritis, Psoriatic"[MeSH]

#6 Search "Psoriasis"[MeSH] OR "Crohn Disease"[MeSH] OR "Colitis, Ulcerative"[MeSH]

#7 Search #5 OR #6

#8 Search #4 AND #7

#9 Search #4 AND #7 Limits: All Adult: 19+ years, English, Publication Date from 1990, Humans

#10 Search #4 AND #7 Limits: All Adult: 19+ years, English, Publication Date from 1990, Review, Humans

#11 Search ("Randomized Controlled Trials"[MeSH] OR "Randomized Controlled Trial"[Publication Type]) OR "Single-Blind Method"[MeSH] OR "Double-Blind Method"[MeSH] OR "Random Allocation"[MeSH]

#12 Search #9 AND #11


#14 Search #9 AND #13

#15 Search adverse OR safety OR withdrawal* OR harm OR mortality OR morbidity OR function* OR toxicity

#16 Search #9 AND #15

#18 Search evaluation studies [pt]

#19 Search #9 AND #18
Appendix B. Studies Already Included in Meta-analyses


Appendix C. Quality Criteria

Assessment of Internal Validity
To assess the internal validity of individual studies, the EPC adopted criteria for assessing the internal validity of individual studies from the US Preventive Services Task Force and the NHS Centre for Reviews and Dissemination.

For Controlled Trials:

Assessment of Internal Validity

1. Was the assignment to the treatment groups really random?
   Adequate approaches to sequence generation:
   - Computer-generated random numbers
   - Random numbers tables
   Inferior approaches to sequence generation:
   - Use of alteration, case record numbers, birth dates or week days
   - Not reported

2. Was the treatment allocation concealed?
   Adequate approaches to concealment of randomization:
   - Centralized or pharmacy-controlled randomization
   - Serially-numbered identical containers
   - On-site computer based system with a randomization sequence that is not readable until allocation
   - Other approaches sequence to clinicians and patients
   Inferior approaches to concealment of randomization:
   - Use of alteration, case record numbers, birth dates or week days
   - Open random numbers lists
   - Serially numbered envelopes (even sealed opaque envelopes can be subject to manipulation)
   - Not reported

3. Were the groups similar at baseline in terms of prognostic factors?

4. Were the eligibility criteria specified?

5. Were outcome assessors blinded to the treatment allocation?

6. Was the care provider blinded?

7. Was the patient kept unaware of the treatment received?

8. Did the article include an intention-to-treat analysis or provide the data needed to calculate it (i.e., number assigned to each group, number of subjects who finished in each group, and their results)?

9. Did the study maintain comparable groups?

10. Did the article report attrition, crossovers, adherence, and contamination?
11. Is there important differential loss to follow-up or overall high loss to follow-up? (Give numbers in each group.)

**Assessment of External Validity (Generalizability)**

1. How similar is the population to the population to whom the intervention would be applied?

2. How many patients were recruited?

3. What were the exclusion criteria for recruitment? (Give numbers excluded at each step.)

4. What was the funding source and role of funder in the study?

5. Did the control group receive the standard of care?

6. What was the length of follow-up? (Give numbers at each stage of attrition.)
Appendix D. Clinical Assessment Scales Commonly Used in Targeted Immune Modulators Trials

**General Health Measures**

**HAQ** - Health Assessment Questionnaire
- HAQ Disability Index (HAQ-DI)
- HAQ visual analog (VAS) pain scale
- VAS patient global health scale
- [http://www.hqlo.com/content/1/1/20](http://www.hqlo.com/content/1/1/20)

**SF-36** - Medical Outcomes Study Short Form 36 Health Survey
- 36 items
- Eight health profiles are derived from summarised scores. All dimensions are independent of each other.
- Scale of 0-100, where higher scores indicate better health and well-being.

**EQ-5D** - EuroQol EQ-5D Quality of Life Questionnaire
- Descriptive system of health-related quality of life states consisting of five dimensions:
  - Mobility
  - Self-care
  - Usual activities
  - Pain/discomfort
  - Anxiety/depression
- Each of which can take one of three responses:
  - No problems
  - Some moderate problems
  - Extreme problems
- [http://www.euroqol.org/web/](http://www.euroqol.org/web/)

**Rheumatoid Arthritis Measures**

**ACR20/50/70** - American College of Rheumatology 20/50/70% improvement
- 20% reductions in tender and swollen joint counts and in at least three of the following: patient’s assessment of pain, patient's global assessment, physician’s global assessment, patient’s assessment of disability, and acute phase reactant (CRP).
- ACR50 and ACR70 were also assessed (defined in a similar manner as ACR20, but with improvement of at least 50% and 70% in the individual measures, respectively).

Example: ACR 50 response

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Endpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tender joints *</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Swollen joints*</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Pain score*</td>
<td>60</td>
<td>20</td>
</tr>
<tr>
<td>Patient’s global activity score*</td>
<td>80</td>
<td>60</td>
</tr>
<tr>
<td>Physician’s global activity score*</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>HAQ-DI</td>
<td>2.0</td>
<td>1.2</td>
</tr>
<tr>
<td>CRP *</td>
<td>3.6</td>
<td>1.4</td>
</tr>
</tbody>
</table>
* at least 50 % improvement
DAS - Disease activity score\(^{197}\)
- Swollen joint count [SJC] and tender joint count [TJC]), employing the 28 joint count; evaluator's and/or patient's global assessment of disease activity (EGA, PGA); and CRP or ESR
- \[\text{DAS28} = (0.56 \times \text{TJC}^{1/2}) + (0.28 \times \text{SJC}^{1/2}) + (0.7 \times \ln \text{[ESR]}) + (0.014 \times \text{PGA [in mm]})\]

**Psoriatic Arthritis Measures**

PsARC - Psoriatic Arthritis Response Criteria\(^{87}\)
- Composite measure requires improvement in two factors (with at least one being a joint score), with worsening in none, of the following four factors: patient and physician global assessments (improvement defined as decrease by $\geq 1$ unit; worsening defined as increase by $\geq 1$ unit); and tender and swollen joint scores
- Improvement defined as decrease by $\geq 30%$; worsening defined as increase by $\geq 30%$.

PASI - Psoriasis area and severity index\(^{88}\)
Composite index of disease severity incorporating measures of;
- Scaling,
  - Erythema, and
  - Induration,
Weighted by severity and affected body surface area

**Ankylosing Spondylitis Measures**

- **BASDAI** - Bath Ankylosing Spondylitis Disease Activity Index\(^{80}\)
- Combined assessment of;
  - Fatigue,
  - Spinal pain,
  - Joint pain,
  - Enthesitis, and
  - Morning stiffness

- **BASFI** - Bath Ankylosing Spondylitis Functional Index\(^{80}\)
- Score ranging from 0 to 10
- Includes 8 questions relating to the patient's function and 2 questions relating to a patient's ability to cope with everyday life.\(^{85}\)

- **BASMI** - Bath Ankylosing Spondylitis Metrology Index\(^{80}\)
- Aggregate score (ranging from 0 to 10) of patient mobility assessments, including tragus-to-wall, lumbar flexion (Schober test), cervical rotation, lumbar side flexion, and intermalleolar distance.

- **ASAS20/50/70** - Assessment in Ankylosing Spondylitis 20% improvement\(^{80}\)
- ASAS20 responder was defined as a patient who showed at least 20% improvement from baseline and had an absolute improvement from baseline of at least 1 unit (on a scale of 0-10) in at least 3 of the following 4 assessment domains:
  - Patient's global assessment,
  - Spinal pain,
  - Function according to the Bath Ankylosing Spondylitis Functional Index (BASFI), and
  - Morning stiffness (the average of the last 2 questions of the BASDAI).
In addition, ASAS20 responders must not have had deterioration from baseline (defined as a worsening of ≥20% and an absolute worsening of at least 1 unit [on a scale of 0-10]) in the potential remaining assessment domain.

40% improvement from baseline and an absolute improvement of at least 2 units [on a scale of 0-10] in at least 3 of the 4 assessment domains defined in the ASAS20 response criteria, with no deterioration from baseline in the potential remaining assessment domain.

**Crohn’s Disease Measures**

**CDAI - Crohn’s Disease Activity Index**

- This index incorporates eight items:
  - Number of liquid or very soft stools
  - Abdominal pain
  - General well-being
  - Extraintestinal manifestations of Crohn's disease
  - Use of opiates to treat diarrhea
  - Abdominal mass
  - Hematocrit
  - Body weight

These yield a composite score ranging from 0 to approximately 600. Higher scores indicate more disease activity; patients with scores of 150 or less are considered to have inactive disease, whereas those with scores above 450 are critically ill.

**CDEIS - Crohn’s Disease Endoscopy Index of Severity**

- Based on the presence of:
  - Deep or superficial ulceration
  - Proportion of ulcerated surface
  - Presence of ulcerated or nonulcerated stenosis in the terminal ileum and four different segments of the colon

**IBDQ – Inflammatory Bowel Disease Questionnaire**

- Scores can range from 32 to 224, and higher scores indicate a better quality of life. It examines the following types of symptoms:
  - Bowel
  - Systemic
  - Emotional
  - Social function

**Juvenile Rheumatoid Arthritis**

**Gianinni’s criteria of improvement**

- 30% improvement from baseline in 3 of any 6 variables in the core set, with no more than 1 of the remaining variables worsening by >30%.
  - Physician global assessment of disease activity;
  - Parent/patient assessment of overall well-being;
  - Functional ability;
  - Number of joints with active arthritis;
  - Number of joints with limited range of motion;
  - Erythrocyte sedimentation rate
Psoriasis Measures

PASI - Psoriasis area and severity index
Composite index of disease severity incorporating measures of
- Scaling
- Erythema
- Induration
Weighted by severity and affected body surface area (body surfaces assessed are head, trunk, and upper and lower limbs)
Score ranges from 0 (no disease) to 72 (maximal disease)

DLQI - Dermatology Life Quality Index
10-item questionnaire that incorporates patients assessment of
- itch
- pain
- feelings of embarrassment and self-consciousness
- problems with their psoriasis treatment
- interference of their psoriasis with their daily activities, relationships, and sexual activity
Scores range from 0 (no impairment) to 30 (maximal impairment)

PSA - Psoriasis Symptom Assessment
16-item measure of 8 psoriasis-related cutaneous symptoms
- hurt
- burning or stinging
- itched
- bothered by water
- irritated
- sensitive
- skin condition bled
- scaling
Contains 2 subscales- one measuring the frequency of the 8 symptoms and the other assessing how troublesome or bothersome the psoriasis symptoms are
# Appendix E. Study Characteristics, Pooled Relative Risks, and Forest Plots of Meta-analyses

## ADALIMUMAB

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Comparisons</th>
<th>Primary outcome</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furst et al. 2003&lt;sup&gt;43&lt;/sup&gt;</td>
<td>RCT</td>
<td>636</td>
<td>24 weeks</td>
<td>ADA +Standard RA therapy / Placebo + Standard RA therapy</td>
<td>safety</td>
<td>Active RA for at least 3 months; DMARD naïve/or on stable regimen; mean disease duration: 10.5 yrs.</td>
</tr>
<tr>
<td>Keystone et al. 2004&lt;sup&gt;44&lt;/sup&gt;</td>
<td>RCT</td>
<td>619</td>
<td>52 weeks</td>
<td>ADA +MTX / Placebo + MTX</td>
<td>Sharp, ACR 20, HAQ</td>
<td>Active RA; on stable MTX regimen; mean disease duration: 11 yrs.</td>
</tr>
<tr>
<td>Van de Putte et al. 2003&lt;sup&gt;45&lt;/sup&gt;</td>
<td>RCT</td>
<td>284</td>
<td>12 weeks</td>
<td>ADA / Placebo</td>
<td>ACR 20</td>
<td>Active RA; had failed at least one DMARD treatment; mean disease duration: 10 yrs.</td>
</tr>
<tr>
<td>Van de Putte et al. 2004&lt;sup&gt;46&lt;/sup&gt;</td>
<td>RCT</td>
<td>544</td>
<td>26 weeks</td>
<td>ADA / Placebo</td>
<td>ACR20</td>
<td>Active RA; had failed at least one DMARD treatment; mean disease duration: 11 yrs.</td>
</tr>
<tr>
<td>Weinblatt et al. 2003&lt;sup&gt;47&lt;/sup&gt;</td>
<td>RCT</td>
<td>271</td>
<td>24 weeks</td>
<td>ADA+MTX / MTX + Placebo</td>
<td>ACR20, HAQ</td>
<td>Active RA;stable MTX regimen; had failed at least one other DMARD; mean disease duration: 12 yrs.</td>
</tr>
</tbody>
</table>

### Relative risk meta-analysis: ACR-20

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Relative risk</th>
<th>95% CI (Koopman)</th>
<th>M-H weight</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.512649</td>
<td>1.262808</td>
<td>1.819429</td>
<td>55.5</td>
</tr>
<tr>
<td>2</td>
<td>2.366746</td>
<td>1.84119</td>
<td>3.091321</td>
<td>32.491115</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>2.48527</td>
<td>10.473312</td>
<td>3.549296</td>
</tr>
<tr>
<td>4</td>
<td>2.234921</td>
<td>1.504395</td>
<td>3.410148</td>
<td>14.104478</td>
</tr>
<tr>
<td>5</td>
<td>4.626866</td>
<td>2.572227</td>
<td>8.746322</td>
<td>4.674419</td>
</tr>
</tbody>
</table>

M-H pooled estimate (Rothman-Boice) of relative risk = 2.100693
Robins-Greenland approximate 95% CI = 1.83305 to 2.407414

Chi-square (for pooled relative risk) = 113.950022 (df = 1) \( P < 0.0001 \)

Q ("non-combinability" for relative risk) = 24.698049 (df = 4) \( P < 0.0001 \)

I²: 83.8%
Relative risk meta-analysis plot (random effects)

Relative risk meta-analysis: ACR-50

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Relative risk</th>
<th>95% CI (Koopman)</th>
<th>M-H weight</th>
<th>M-H pooled estimate (Rothman-Boice) of relative risk = 3.536893</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.552833</td>
<td>1.80314</td>
<td>3.63624</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>4.17033</td>
<td>2.711696</td>
<td>6.522056</td>
<td>12.861066</td>
</tr>
<tr>
<td>3</td>
<td>16.527778</td>
<td>2.954667</td>
<td>96.371191</td>
<td>0.507042</td>
</tr>
<tr>
<td>4</td>
<td>2.607407</td>
<td>1.365527</td>
<td>5.10824</td>
<td>6.044776</td>
</tr>
<tr>
<td>5</td>
<td>6.847761</td>
<td>3.047254</td>
<td>16.177401</td>
<td>2.596899</td>
</tr>
</tbody>
</table>

Robins-Greenland approximate 95% CI = 2.774584 to 4.508643

Chi-square (for pooled relative risk) = 104.031248 (df = 1) $P < 0.0001$

Q ("non-combinability" for relative risk) = 9.132299 (df = 4) $P = 0.0579$

$I^2 = 56.2\%$
Relative risk meta-analysis: ACR-70

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Relative risk</th>
<th>95% CI (Koopman)</th>
<th>M-H weight</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.278545</td>
<td>2.294726</td>
<td>8.036822</td>
<td>5.5</td>
</tr>
<tr>
<td>2</td>
<td>4.879342</td>
<td>2.568811</td>
<td>9.421447</td>
<td>6.092084</td>
</tr>
<tr>
<td>3</td>
<td>16.531034</td>
<td>1.715513</td>
<td>164.871224</td>
<td>0.253497</td>
</tr>
<tr>
<td>4</td>
<td>6.111111</td>
<td>1.66042</td>
<td>23.11434</td>
<td>1.343284</td>
</tr>
<tr>
<td>5</td>
<td>5.552239</td>
<td>1.873092</td>
<td>17.136578</td>
<td>1.55814</td>
</tr>
</tbody>
</table>

M-H pooled estimate (Rothman-Boice) of relative risk = 5.038857
Robins-Greenland approximate 95% CI = 3.353377 to 7.571496

Chi-square (for pooled relative risk) = 60.586043 (df = 1) \( P < 0.0001 \)

Q (“non-combinability” for relative risk) = 1.034209 (df = 4) \( P = 0.9046 \)

\( I^2 : 0\% \)
Relative risk meta-analysis plot (random effects)

- **Furst 2003**: 4.28 (2.29, 8.04)
- **Keystone 2004**: 4.88 (2.57, 9.42)
- **Van de Putte 2003**: 16.53 (1.72, 164.87)
- **Van de Putte 2004**: 6.11 (1.66, 23.11)
- **Weinblatt 2003**: 5.55 (1.87, 17.14)
- **combined [random]**: 4.91 (3.27, 7.36)

**Relative risk (95% confidence interval)**

**Relative risk meta-analysis plot (random effects)**
ANAKINRA

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Comparisons</th>
<th>Primary outcome</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bresnihan et al. 1998[^2]</td>
<td>RCT</td>
<td>472</td>
<td>24 weeks</td>
<td>AKA / Placebo</td>
<td>ACR-N</td>
<td>&gt; 6 months active RA &lt; 8 years; mean disease duration: 3.7-4.3 years</td>
</tr>
<tr>
<td>Cohen et al. 2002[^3]</td>
<td>RCT</td>
<td>419</td>
<td>24 weeks</td>
<td>AKA+MTX / MTX+ Placebo</td>
<td>ACR 20</td>
<td>&gt; 6 months active RA &lt; 12 years; stable MTX regimen; mean disease duration: 6.3-8.8 years</td>
</tr>
</tbody>
</table>

Relative risk meta-analysis: ACR-20

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Relative risk</th>
<th>95% CI (Koopman)</th>
<th>M-H weight</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.450566</td>
<td>1.045564</td>
<td>2.052383</td>
<td>21.031161</td>
</tr>
<tr>
<td>2</td>
<td>2.619469</td>
<td>1.491026</td>
<td>4.769021</td>
<td>6.647059</td>
</tr>
<tr>
<td>3</td>
<td>1.734182</td>
<td>1.312326</td>
<td>2.304111</td>
<td>27.44511</td>
</tr>
</tbody>
</table>

M-H pooled estimate (Rothman-Boice) of relative risk = 1.732727
Robins-Greenland approximate 95% CI = 1.413511 to 2.12403

Chi-square (for pooled relative risk) = 27.996519 (df = 1) P < 0.0001

Q ("non-combinability" for relative risk) = 2.927509 (df = 2) P = 0.2314

I²: 31.68%
Relative risk meta-analysis: ACR-50

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Relative risk</th>
<th>95% CI (Koopman)</th>
<th>M-H weight</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.825431</td>
<td>0.958312</td>
<td>3.546318</td>
<td>6.572238</td>
</tr>
<tr>
<td>2</td>
<td>6.548673</td>
<td>1.790818</td>
<td>24.879122</td>
<td>1.208556</td>
</tr>
<tr>
<td>3</td>
<td>2.1586</td>
<td>1.318936</td>
<td>3.55346</td>
<td>9.98004</td>
</tr>
</tbody>
</table>

M-H pooled estimate (Rothman-Boice) of relative risk = 2.334041
Robins-Greenland approximate 95% CI = 1.590173 to 3.425885

Chi-square (for pooled relative risk) = 18.739732 (df = 1)  $P < 0.0001$

Q ("non-combinability" for relative risk) = 2.631496 (df = 2)  $P = 0.2683$

$\hat{\tau}^2: 23.99\%$
Relative risk meta-analysis: ACR-70

| Stratum | Relative risk | 95% CI (Koopman) | M-H weight | M-H weight
|---------|---------------|------------------|------------|----------------
| 1       | 1.043103      | 0.138162         | 7.92919    | Bresnihan 1998
| 2       | 9.230088      | 0.942796         | 93.142286  | Cohen 2002
| 3       | 3.012         | 1.158293         | 7.883807   | Cohen 2004

M-H pooled estimate (Rothman-Boice) of relative risk = 3.179859
Robins-Greenland approximate 95% CI = 1.345937 to 7.512612

Chi-square (for pooled relative risk) = 6.955041 (df = 1) \( P = 0.0084 \)

Q ("non-combinability" for relative risk) = 1.382147 (df = 2) \( P = 0.501 \)

\( I^2 : 0\% \)
## ETANERCEPT

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Comparisons</th>
<th>Primary outcome</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Klareskog et al. 2004</td>
<td>RCT</td>
<td>682</td>
<td>52 weeks</td>
<td>ETA / MTX / MTX + ETA</td>
<td>Sharp</td>
<td>&gt; 6 months active RA; ACR functional class I-III; unsatisfactory response to at least one DMARD other than MTX; mean disease duration: 6.5 yrs.</td>
</tr>
<tr>
<td>Lan et al. 2004</td>
<td>RCT</td>
<td>58</td>
<td>12 weeks</td>
<td>ETA+ MTX / Placebo + MTX</td>
<td>Number of swollen/tender joints</td>
<td>Active RA &gt; one year; stable MTX for 4 weeks; mean disease duration: NR</td>
</tr>
<tr>
<td>Moreland et al. 1997</td>
<td>RCT</td>
<td>180</td>
<td>12 weeks</td>
<td>ETA / Placebo</td>
<td>Number of swollen/tender joints</td>
<td>Active RA; failed 1 to 4 DMARD treatments; mean disease duration: NR</td>
</tr>
<tr>
<td>Moreland et al. 1999</td>
<td>RCT</td>
<td>234</td>
<td>12 weeks</td>
<td>ETA / Placebo</td>
<td>ACR20/50</td>
<td>Active RA; failed 1 to 4 DMARD treatments other than MTX; mean disease duration: 12 yrs.</td>
</tr>
<tr>
<td>Weinblatt et al. 1999</td>
<td>RCT</td>
<td>89</td>
<td>24 weeks</td>
<td>ETA+ MTX / Placebo + MTX</td>
<td>ACR 20</td>
<td>Active RA; &gt; 6 months MTX, stable &gt;1 month; mean disease duration: 13 years</td>
</tr>
</tbody>
</table>

### Relative risk meta-analysis: ACR-20

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Relative risk</th>
<th>95% CI (Koopman)</th>
<th>M-H weight</th>
<th>M-H pooled estimate (Rothman-Boice) of relative risk = 1.83981</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.264839</td>
<td>1.111763</td>
<td>1.447291</td>
<td>67.941176 Klareskog 2004</td>
</tr>
<tr>
<td>2</td>
<td>2.6</td>
<td>1.649044</td>
<td>4.544377</td>
<td>5 Lan 2004</td>
</tr>
<tr>
<td>3</td>
<td>5.501166</td>
<td>3.234162</td>
<td>9.749303</td>
<td>5.43038 Moreland 1999</td>
</tr>
<tr>
<td>4</td>
<td>5.5</td>
<td>2.730932</td>
<td>11.900985</td>
<td>3 Moreland 1997</td>
</tr>
<tr>
<td>5</td>
<td>2.669492</td>
<td>1.547005</td>
<td>5.107559</td>
<td>5.303371 Weinblatt 1999</td>
</tr>
</tbody>
</table>

M-H pooled estimate (Rothman-Boice) of relative risk = 1.83981
Robins-Greenland approximate 95% CI = 1.618818 to 2.09097

Chi-square (for pooled relative risk) = 87.193615 (df = 1) $P < 0.0001$

$Q$ ("non-combinability" for relative risk) = 56.969838 (df = 4) $P < 0.0001$

$I^2$: 92%
Relative risk meta-analysis plot (random effects)

Relative risk meta-analysis: ACR-50

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Relative risk</th>
<th>95% CI (Koopman)</th>
<th>M-H weight</th>
<th>M-H pooled estimate (Rothman-Boice) of relative risk = 2.585038</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.757365</td>
<td>1.446</td>
<td>2.153791</td>
<td>41.267974 (Klareskog 2004)</td>
</tr>
<tr>
<td>2</td>
<td>6.333333</td>
<td>2.362599</td>
<td>18.757771</td>
<td>1.5 (Lan 2004)</td>
</tr>
<tr>
<td>3</td>
<td>8.205128</td>
<td>3.598388</td>
<td>19.451313</td>
<td>2.468354 (Moreland 1999)</td>
</tr>
<tr>
<td>4</td>
<td>8.333333</td>
<td>2.998444</td>
<td>24.815338</td>
<td>1.5 (Moreland 1997)</td>
</tr>
<tr>
<td>5</td>
<td>11.694915</td>
<td>2.26005</td>
<td>67.188802</td>
<td>0.662921 (Weinblatt 1999)</td>
</tr>
</tbody>
</table>

Chi-square (for pooled relative risk) = 92.446788 (df = 1) \( P < 0.0001 \)

Q ("non-combinability" for relative risk) = 30.10553 (df = 4) \( P < 0.0001 \)

\( I^2: 87\% \)
Relative risk meta-analysis plot (random effects)

Relative risk meta-analysis: ACR-70

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Relative risk</th>
<th>95% CI (Koopman)</th>
<th>M-H weight</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.328338</td>
<td>1.689058</td>
<td>3.237337</td>
<td>Klareskog 2004</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>1.635418</td>
<td>149.135742</td>
<td>Lan 2004</td>
</tr>
<tr>
<td>3</td>
<td>15.384615</td>
<td>2.714878</td>
<td>90.264012</td>
<td>Moreland 1999</td>
</tr>
<tr>
<td>4</td>
<td>9.661017</td>
<td>1.061662</td>
<td>95.694514</td>
<td>Weinblatt 1999</td>
</tr>
</tbody>
</table>

M-H pooled estimate (Rothman-Boice) of relative risk = 2.910097
Robins-Greenland approximate 95% CI = 2.116173 to 4.001877

Chi-square (for pooled relative risk) = 43.187838 (df = 1) \( P < 0.0001 \)

Q ("non-combinability" for relative risk) = 6.455625 (df = 3) \( P = 0.0914 \)

\( I^2: 53\% \)
INFLIXIMAB

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study design</th>
<th>N</th>
<th>Duration</th>
<th>Comparisons</th>
<th>Primary outcome</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abe et al., 2006&lt;sup&gt;51&lt;/sup&gt;</td>
<td>RCT</td>
<td>147</td>
<td>14 weeks</td>
<td>INF+ MTX / Placebo + MTX</td>
<td>ACR 20</td>
<td>&gt; 6 months history of active RA; mean disease duration 7.9 yrs.</td>
</tr>
<tr>
<td>Kavanaugh et al. 2000&lt;sup&gt;200&lt;/sup&gt;</td>
<td>RCT</td>
<td>28</td>
<td>12 weeks</td>
<td>INF+ MTX / Placebo + MTX</td>
<td>ACR 20</td>
<td>RA &lt; 15 years; MTX &gt; 3 months; mean disease duration 4.9 – 7.5 years</td>
</tr>
<tr>
<td>Maini et al. 1998&lt;sup&gt;71&lt;/sup&gt;</td>
<td>RCT</td>
<td>43</td>
<td>26 weeks</td>
<td>INF+ MTX / Placebo + MTX</td>
<td>Paulus 20</td>
<td>MTX &gt; 6 months; mean disease duration 7.6 – 114.3 years</td>
</tr>
<tr>
<td>Maini et al. 1999&lt;sup&gt;48&lt;/sup&gt;</td>
<td>RCT</td>
<td>428</td>
<td>30 weeks</td>
<td>INF+MTX / Placebo + MTX</td>
<td>ACR 20</td>
<td>MTX stable &gt; 4 weeks; mean disease duration 7.2 – 9.0 years</td>
</tr>
<tr>
<td>Westhovens et al., 2006&lt;sup&gt;57&lt;/sup&gt;</td>
<td>RCT</td>
<td>1084</td>
<td>22 weeks</td>
<td>INF+ MTX / Placebo + MTX</td>
<td>ACR 20</td>
<td>Active RA despite MTX treatment; median disease duration: 15 yrs</td>
</tr>
</tbody>
</table>

Relative risk meta-analysis: ACR-20, St. Clair et al. removed
### Relative risk meta-analysis: ACR-50, St. Clair et al. removed

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Relative risk</th>
<th>95% CI (Koopman)</th>
<th>% Weights (fixed, random)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.8775</td>
<td>1.576166 10.168522</td>
<td>7.395122 7.617304 Abe 2006</td>
</tr>
<tr>
<td>2</td>
<td>1.5</td>
<td>0.269401  9.804675</td>
<td>1.811805 1.695111 Kavanough 2000</td>
</tr>
<tr>
<td>5</td>
<td>13.034483</td>
<td>1.645997 126.445188</td>
<td>0.916436 0.963797 Maini 1998</td>
</tr>
<tr>
<td>6</td>
<td>3.493759</td>
<td>2.497169  4.931648</td>
<td>59.651669 62.136605 Westhovens 2006</td>
</tr>
</tbody>
</table>
Fixed effects (Mantel-Haenszel, Rothman-Boice)
Pooled relative risk = 3.763542 (95% CI = 2.870766 to 4.933963)
Chi² (test relative risk differs from 1) = 92.029162  (df = 1)  P < 0.0001

Non-combinability of studies
Cochran Q = 1.846138  (df = 5)  P = 0.87
Moment-based estimate of between studies variance = 0
I² (inconsistency) = 0% (95% CI = 0% to 61%)

Relative risk meta-analysis plot (random effects)

- Abe 2006  3.88 (1.58, 10.17)
- Kavanaugh 2000  1.50 (0.27, 9.80)
- Lipsky 2000  4.14 (2.09, 8.56)
- Maini 1999  4.10 (2.07, 8.48)
- Maini 1998  13.03 (1.65, 126.45)
- Westhovens 2006  3.49 (2.50, 4.93)
- combined [random]  3.68 (2.81, 4.82)
### Relative risk meta-analysis: ACR-20

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Relative risk</th>
<th>95% CI (Koopman)</th>
<th>% Weights (fixed, random)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.435455</td>
<td>1.475729 4.287878</td>
<td>2.556272 7.081826</td>
</tr>
<tr>
<td>2</td>
<td>1.512649</td>
<td>1.262808 1.819429</td>
<td>18.9594 8.504502</td>
</tr>
<tr>
<td>3</td>
<td>3.5</td>
<td>0.804588 20.402506</td>
<td>0.227741 2.299906</td>
</tr>
<tr>
<td>4</td>
<td>2.366746</td>
<td>1.84119 3.091321</td>
<td>11.099318 8.286448</td>
</tr>
<tr>
<td>5</td>
<td>1.131313</td>
<td>1.032425 1.245886</td>
<td>29.398629 8.668226</td>
</tr>
<tr>
<td>6</td>
<td>2.6</td>
<td>1.649044 4.544377</td>
<td>1.708054 7.217192</td>
</tr>
<tr>
<td>7</td>
<td>7.241379</td>
<td>1.545446 41.553324</td>
<td>0.230389 2.242662</td>
</tr>
<tr>
<td>8</td>
<td>3.036863</td>
<td>1.947037 4.928533</td>
<td>4.070597 7.428933</td>
</tr>
<tr>
<td>9</td>
<td>5.5</td>
<td>2.730932 11.900985</td>
<td>1.024833 5.993716</td>
</tr>
<tr>
<td>10</td>
<td>5.242165</td>
<td>2.849323 10.009865</td>
<td>1.51779 6.593108</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
<td>2.48527 10.473312</td>
<td>1.212478 6.108029</td>
</tr>
<tr>
<td>12</td>
<td>2.234921</td>
<td>1.504395 3.410148</td>
<td>4.818243 7.695944</td>
</tr>
<tr>
<td>13</td>
<td>2.669492</td>
<td>1.547005 5.107559</td>
<td>1.811689 6.730877</td>
</tr>
<tr>
<td>14</td>
<td>4.626866</td>
<td>2.572227 8.746322</td>
<td>1.596832 6.672768</td>
</tr>
<tr>
<td>15</td>
<td>2.337941</td>
<td>1.934503 2.850269</td>
<td>19.767735 8.475861</td>
</tr>
</tbody>
</table>

**Fixed effects (Mantel-Haenszel, Rothman-Boice)**
- Pooled relative risk = 2.025654 (95% CI = 1.86971 to 2.194605)
- Chi² (test relative risk differs from 1) = 298.267679 (df = 1) \( P < 0.0001 \)

**Non-combinability of studies**
- Cochran Q = 199.671473 (df = 14) \( P < 0.0001 \)
- Moment-based estimate of between studies variance = 0.33231
- \( I^2 \) (inconsistency) = 93% (95% CI = 90.5% to 94.5%)
Relative risk meta-analysis: ACR-50

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Relative risk</th>
<th>95% CI (Koopman)</th>
<th>% Weights (fixed, random)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.8775</td>
<td>1.576166 to 10.168522</td>
<td>2.156127 to 6.566185 Abe 2006</td>
</tr>
<tr>
<td>2</td>
<td>2.552833</td>
<td>1.80314 to 3.63624</td>
<td>14.262778 to 10.290635 Furst 2003</td>
</tr>
<tr>
<td>3</td>
<td>1.5</td>
<td>0.269401 to 9.804675</td>
<td>0.528251 to 2.67639 Kavanough 2000</td>
</tr>
<tr>
<td>5</td>
<td>1.601378</td>
<td>1.352304 to 1.911821</td>
<td>39.080219 to 10.998743 Klareskog 2004</td>
</tr>
<tr>
<td>6</td>
<td>6.333333</td>
<td>2.362599 to 18.757771</td>
<td>1.188565 to 5.896628 Lan 2004</td>
</tr>
<tr>
<td>7</td>
<td>13.034483</td>
<td>1.645997 to 126.445188</td>
<td>0.267197 to 1.695889 Maini 1998</td>
</tr>
<tr>
<td>8</td>
<td>4.104202</td>
<td>2.066097 to 8.480455</td>
<td>4.406206 to 8.068578 Maini 1999</td>
</tr>
<tr>
<td>9</td>
<td>8.333333</td>
<td>2.984444 to 24.815338</td>
<td>1.188565 to 5.801936 Moreland 1997</td>
</tr>
<tr>
<td>10</td>
<td>7.948718</td>
<td>3.130217 to 20.937153</td>
<td>1.564693 to 6.480541 Moreland 1999</td>
</tr>
<tr>
<td>11</td>
<td>16.527778</td>
<td>2.954667 to 96.371194</td>
<td>0.401768 to 2.847801 Van de Putte 2003</td>
</tr>
<tr>
<td>12</td>
<td>2.607407</td>
<td>1.365257 to 5.10824</td>
<td>4.789739 to 8.39877 Van de Putte 2004</td>
</tr>
<tr>
<td>13</td>
<td>11.694915</td>
<td>2.26005 to 67.188805</td>
<td>0.525283 to 2.927232 Weinblatt 1999</td>
</tr>
<tr>
<td>14</td>
<td>6.847761</td>
<td>3.047254 to 16.177401</td>
<td>2.057722 to 7.204715 Weinblatt 2003</td>
</tr>
<tr>
<td>15</td>
<td>3.493759</td>
<td>2.497169 to 4.931648</td>
<td>17.39208 to 10.34049 Westhovens 2006</td>
</tr>
</tbody>
</table>

Fixed effects (Mantel-Haenszel, Rothman-Boice)
Pooled relative risk = 3.022076 (95% CI = 2.646599 to 3.450823)
Chi² (test relative risk differs from 1) = 266.947484 (df = 1) P < 0.0001
Non-combinability of studies
Cochran Q = 74.645572 (df = 14)  P < 0.0001
Moment-based estimate of between studies variance = 0.349595
I² (inconsistency) = 81.2% (95% CI = 69.1% to 87.2%)
Appendix F. Abstract-only Studies (Not Included)


12. Ganguly R. Etanercept therapy provides clinically meaningful improvement in dermatology quality of life index in patients with chronic plaque psoriasis Abstract P103 European Congress on


36. van Vollenhoven R, Breedveld FC, Kavanaugh AF, Cohen SB, Perez JL, Spencer-Green GT. The Clinical and Radiographic Efficacy of Every-Other-Week vs. Weekly Dosing Frequency of


40. Weisman MH, Strand V, Cifaldi MA, Sterz R. Adalimumab (HUMIRA®) Plus Methotrexate is Superior to MTX Alone in Improving Physical Function, as Measured by the SF-36, in Patients with Early Rheumatoid Arthritis. OASIS Online Abstract Submission and Invitation System 2005;Presentation 1018(Poster Board 320).

## Appendix G. Characteristics of Studies with Poor Internal Validity

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Sample Size</th>
<th>Intervention</th>
<th>Reason for Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bathon et al.</td>
<td>Pooled data analysis</td>
<td>2,402</td>
<td>Etanercept</td>
<td>Bias</td>
</tr>
<tr>
<td>Gerloni et al.</td>
<td>Open label prospective trial</td>
<td>24</td>
<td>Infliximab</td>
<td>High LTF</td>
</tr>
<tr>
<td>Moreland et al.</td>
<td>Pooled retrospective analysis</td>
<td>714</td>
<td>Etanercept</td>
<td>High LTF; no ITT analysis</td>
</tr>
</tbody>
</table>

ITT: intention to treat; LTF: loss to follow-up
Appendix H. Acknowledgements

Acknowledgements

Reviewers

We gratefully acknowledge the following individuals who reviewed the initial draft of this report and provided us with valuable and constructive feedback.

Stanley Cohen, MD
Clinical Professor of Internal Medicine
University of Texas
Southwestern Medical School at Dallas
Medical Director
Radiant Research
Dallas, Texas

Roy Fleischmann, MD
Clinical Professor of Medicine
University of Texas
Southwestern Medical Center at Dallas

Marian S. McDonagh, PharmD
Assistant Professor
Department of Medical Informatics and Clinical Epidemiology
Oregon Health and Sciences University
School of Medicine
References


