Relative versus absolute goals of therapies for RA: ACR 20 or ACR 50 responses versus target values for "near remission" of DAS or single measures

T. Pincus¹, T. Sokka¹², A. Kavanaugh³

ABSTRACT
Therapies for rheumatoid arthritis (RA) may be assessed according to relative levels of measures to compare efficacy to another therapy or to a placebo, as in the American College of Rheumatology (ACR) 20%, 50%, or 70% (ACR 20 ACR 50 and ACR 70) responses, or by absolute levels of measures, as in disease activity scores (DAS), ACR criteria for remission, or "target values" of specific measures. Regulatory considerations have emphasized primarily relative comparisons to a placebo or standard therapy, derived in part from the weak efficacy of traditional disease modifying anti-rheumatic drugs (DMARDs). While improvement compared to placebo certainly indicates efficacy, it is of concern that measures of inflammatory activity, such as swollen joints and the erythrocyte sedimentation rate (ESR), may be stable or improved over periods of 5-10 years, while measures of damage, such as joint deformity and radiographic changes, may progress over the same period in the same patients. These findings suggest that improvement at a level of 20% or 50% may deter but not prevent severe long-term outcomes of radiographic progression, functional declines, work disability, and premature mortality, seen in most patients until the middle 1990s. Outcomes appear to be improved at this time, associated with aggressive treatment strategies and more powerful therapies, including biologic agents. In the Finnish Rheumatoid Arthritis Combination Therapy Trial (FinRACo), no patient who was in remission after 6 months was receiving work disability payments 4 1/2 years later, compared to 22% of patients who had ACR 20 or 50 responses and 54% of patients who did not have ACR 20 responses after 6 months who were all receiving work disability payments after 5 years. These findings suggest that absolute targets, including remission, may be realistic contemporary goals, with aggressive treatment strategies and more effective DMARDs and biologic agents.

Introduction
Therapies for rheumatoid arthritis (RA) are evaluated in formal research studies using two approaches. The first involves efficacy relative to another therapy or to a placebo, as in the American College of Rheumatology (ACR) 20%, 50%, or 70% (ACR 20 ACR 50 and ACR 70) responses (1, 2). The second approach involves absolute efficacy, such as a disease activity score (DAS) (3, 4), ACR criteria for remission (5), or "target values" of specific measures (6). The process of regulatory approval has emphasized primarily relative approaches, in which approval of new therapies is based on demonstration of statistically-significant differences compared to a placebo or standard therapy, according to a pre-defined endpoint (7). This approach to policy may be derived in part from the fact that traditional disease modifying anti-rheumatic drugs (DMARDs) had relatively weak efficacy, and absolute targets were unrealistic.

Development of powerful DMARDs such as methotrexate (8-10), cyclosporine A (11), leflunomide (12, 13), biologic agents to inhibit tumor necrosis factor alpha (TNFα) [etanercept (14, 15), infliximab (16, 17), and adalimumab (18, 19)], and the interleukin-1 receptor antagonist, anakinra (20, 21), has provided major advances in RA over the last two decades. The efficacy of these therapies in clinical trials has been documented primarily as relative differences in ACR 20 or 50 responses with active treatment versus placebo or, in partial responders to methotrexate, the agent in combination with methotrexate versus methotrexate monotherapy. Little information is available concern-
ing possible remission in these trials. However, relatively few patients appear likely to have entered into remission, in part because patients selected for these trials had relatively severe clinical status, as only a small fraction of consecutive patients seen in one setting met the inclusion criteria for most recent clinical trials (22, 23).

Further discussion of this subject requires recognition of differences between measures of inflammatory activity, measures of damage, and outcome measures (Table I). The measures included in the ACR Core Data Set (1, 24, 25) and the DAS (26, 27) may be classified broadly as short-term measures primarily of disease activity – i.e., joint swelling, joint tenderness and ESR or CRP, or as measures of activity and damage – i.e., functional disability, pain, and patient and physician and global assessment. Activity measures are sensitive to change over weeks to months, and are regarded as short-term surrogate markers for measures of long-term joint damage, such as joint deformity and radiographic progression, and clinical outcomes, such as joint replacement surgery and premature mortality (28), which develop over years to decades. One measure of damage, the radiographic score, is included in the Core Data Set in clinical trials conducted over one year or longer (Table I). No other measures of damage or long-term outcomes are included in most clinical trials. Suppression of inflammation at a level of 20% or 50%, i.e., ACR 20 or ACR 50, appears unlikely to provide optimal improvement for patients. Furthermore, measures of inflammatory activity may be stable or improved over periods of 5-10 years while measures of damage may progress (Table II). A summary of these reports was presented in a previous article (29), but since the information does not appear to be widely known in the rheumatology community, it is summarized below. A brief overview of the values for the Core Data Set measures at the conclusion of pivotal trials of biologic agents is then presented, followed by a discussion of the possible advantages of absolute values or remission as goals of contemporary RA clinical trials.

**Reports concerning improvement in inflammatory activity and simultaneous progression of damage over 5-10 years**

Over the last decade, as noted, several reports have documented that measures of inflammatory activity such as those included in the Core Data set or DAS were improved over 5-10 years, while radiographic progression and in some cases physical function showed disease progression in the same patients over the same period (Table II). One of the earliest studies which called attention to this phenomenon was by Scott in 1984 (30), who reported improvements in ESR over one year in 75% of patients, while radiographic progression was also seen in almost all of the same patients (Fig. 1). Further reports by Scott indicated similar findings over longer periods (31, 32).

Another study reported in 1984 indicated that morning stiffness was improved in half of patients over 9 years, while more than 90% of patients experienced progression of functional disability (33). Furthermore, grip strength and walking time also showed progression of damage in these patients.

Hawley and Wolfe (34) found that joint tenderness, morning stiffness and ESR, hemoglobin, grip strength and global severity were improved, unchanged or slightly worse in 157 patients over 2, 5 and 10 years. However, health assessment questionnaire (HAQ) disability scores showed evidence of substantial disease progression with an effect size 2.5-fold greater than any other clinical measure studied. In other reports, which included many of the same patients, there was also evidence of radiographic progression in most patients (35).

Egsnose *et al.* (36) reported improvement in measures of disease activity over 5 years, including the number of swollen joints, morning stiffness, the Ritchie Articular Index for joint tenderness, and grip strength, with contempo-

---

**Table I. Measures of disease ACTIVITY, joint DAMAGE and OUTCOMES in rheumatoid arthritis.**

<table>
<thead>
<tr>
<th>Type of measure</th>
<th>Disease ACTIVITY measures included in most clinical trials</th>
<th>Measures of disease ACTIVITY and/or joint DAMAGE included in most clinical trials</th>
<th>Joint DAMAGE markers and long-term OUTCOMES not included in most clinical trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joint count</td>
<td>Tenderness or pain on motion (C)</td>
<td></td>
<td>Joint with limited motion or deformity;</td>
</tr>
<tr>
<td></td>
<td>Swelling (C)</td>
<td></td>
<td>Joint replacement surgery</td>
</tr>
<tr>
<td>Laboratory</td>
<td>Acute phase reactant –</td>
<td></td>
<td>Radiographic damage (C*);</td>
</tr>
<tr>
<td></td>
<td>ESR or CRP (C)</td>
<td></td>
<td>Joint replacement surgery</td>
</tr>
<tr>
<td>Radiographic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient questionnaire</td>
<td>Pain (C)</td>
<td></td>
<td>Functional disability (C)</td>
</tr>
<tr>
<td>Global</td>
<td>American Rheumatism Association (ARA) functional class</td>
<td></td>
<td>Work disability</td>
</tr>
<tr>
<td></td>
<td>Physician assessment of global status (C)</td>
<td></td>
<td>Comorbid diseases</td>
</tr>
<tr>
<td></td>
<td>Patient assessment of global status (C)</td>
<td></td>
<td>Extra-articular disease</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Premature mortality</td>
</tr>
</tbody>
</table>

(C) Included in Core Data Set (1, 24, 25) recommended for use in clinical trials. *Included in Core Data Set for studies longer than one year.
raneous radiographic progression in a cohort of 75 patients who had RA for 2 years or less duration at baseline. Improvement in activity measures was less in patients in whom the institution of anti-rheumatic therapy had been delayed, compared to patients who received early treatment with auranofin. Furthermore, radiographic progression was significantly less in patients who had received early treatment than in those with delayed treatment. HAQ disability scores were improved over 5 years in patients treated with early DMARDs, but were similar to baseline in those who had delayed DMARD therapy (36).

Fex et al. (37) examined changes over 5 years in 113 patients who had a mean disease duration of 11.4 months at baseline and were monitored prospectively for 5 years. Values for morning stiffness, pain, general health, Ritchie index, HAQ scores, ESR and hemoglobin were similar or improved from baseline after 5 years of observation. However, radiographic scores indicated significant progression in these patients.

Mulherin et al. (38) reported significant improvement in grip strength, the Ritchie articular index, hemoglobin and ESR, and non-significant improvement in pain scores and morning stiffness, while there was significant progression of radiographic changes in 40 patients over 6 years (Fig. 2). They pointed out that "articular erosion continues in RA despite clinical improvement, and is accelerated in those with evidence of continuing synovial inflammation, reflected in the clinical and laboratory measures of disease activity."

Callahan et al. (39) reported that joint tenderness, swelling, ESR, hemoglobin, morning stiffness, pain, and the modified health assessment questionnaire (MHAQ) results were unchanged or improved over 5 years in 100 patients, while scores for radiographic damage as well as joint deformity, grip strength and walking time indicated disease progression (Fig. 3). The effect size of scores for joint swelling indicated improvement at the 20% level, suggesting that this level of improvement may not necessarily be associated with the prevention of radiographic progression.

Leirisalo-Repo et al. (40) reported stable or improved swollen joint counts...
and tender joint counts, with progression of the Larsen scores at 2-13 years after baseline in 145 patients with RA who had disease for only a mean of 8 months at baseline. Furthermore, HAQ scores were improved at years 2 and 3 in this group, possibly reflecting improvement in the component of the HAQ sensitive to disease activity. HAQ scores after 8 years indicated progression of disability.

Graudal et al. (41) found improvement in joint swelling, ESR, and hemoglobin, while radiographic progression was seen in most of 112 patients with RA over 4-22 years. There was some relationship between the severity of joint swelling and a high ESR and low hemoglobin and subsequent radiographic progression, while the association of joint tenderness with radiographic progression was weak. Radiographic damage was greater in patients with greater inflammatory activity over time.

Welsing et al. (42) found that disease activity according to the DAS remained more or less the same over 9 years in patients with early RA. However, functional capacity worsened after an initial improvement, and radiographic scores worsened over the 9-year observation period.

These studies indicate progression of radiographic damage and decline of physical function in most studies, with significant improvements in measures of inflammatory activity included in the Core Data Set and DAS. The data serve as an important caution concerning the anticipated long-term outcomes of patients showing improvement in clinical trials. It is difficult to extrapolate the extent of clinical improvement from the reports, but many indicated a range of at least 20% or higher, while progression of radiographic damage and further functional declines are seen. Taken together, the data raise a question as to whether the goal of future clinical trials might be levels higher than ACR 20, even higher than ACR 50, or absolute measures such as DAS (3) or ACR criteria for remission (5), or "target values" of specific measures (6). Further concerns about ACR 20 as an endpoint in RA trials are presented in the next section.

**End points in clinical trials of new biological agents**

As noted above, driven in part by regulatory considerations, clinical trials of new biological agents to inhibit TNFα, etanercept (14, 15), infliximab (16, 17), and adalimumab (18, 19) and the interleukin-1 receptor antagonist, anakinra (20, 21), have been conducted with primary endpoints as relative differences in ACR 20 or 50 responses. Furthermore, inclusion criteria for these trials selected for patients who had a relatively severe clinical status, as only a small fraction of consecutive patients seen in one setting met the inclusion criteria (22, 23). This selection is appropriate for early trials of new agents. However, selection limits trials to relatively few patients who would enter into remission, and such data have not been commonly reported.
Indeed, the status of many patients at the conclusion of pivotal clinical trials was not only near remission, but also reflected significant disease activity (14, 18, 43, 44) (Table III). For example, the mean swollen joint count (total 68) after 6 months of therapy was 13 for etanercept therapy, 9 for infliximab, 11.3 for anakinra, and 6.9 for adalimumab. Mean pain scores (0-100) were 32, 38, 34 and 28 for patients at the conclusion of these 4 trials. These levels reflect improvements of 31-61%, indicating a powerful reduction of inflammatory indicators, but reflecting the severe status of the patients who were eligible to be included in these trials.

These data emphasize two important points. First, most of the data concerning the efficacy of new therapies for RA, including biological anti-TNF therapies, are available only in selected patients who have very severe disease (22, 23). Second, the results of pivotal clinical trials document statistically significant differences between active versus placebo treatment, but do not necessarily reflect optimal status for patients with RA. The continued presence of active inflammation is particularly worrisome, as markers of inflammation may improve over 5 years or more, while markers of damage to joints may show continued progression, as discussed above.

### Absolute treatment goals, including DAS, target values, and remission

Rheumatologists have spoken of "remission" and "remission-inducing therapy" in RA for many years (5), much as oncologists speak of "no evidence of disease" in patients with neoplastic disease (45). However, sustained remissions in RA were unusual with traditional DMARDs (46) — many patients who appeared to enter remission likely had a self-limited inflammatory arthritis (47). Patients with RA seen in clinical settings, even in the community (48), generally have persistent inflammatory symmetrical arthritis (PISA) (49, 50), or Type III RA (47), in which apparent remissions in clinical care of RA may be temporary and followed by exacerbations and long-term disease progression (46, 51, 52). Furthermore, as the fundamental dysregulation in RA remains unknown (as in other dysregulatory conditions such as hypertension and diabetes), the term "remission" in RA includes continuing therapy, and not a drug-free remission.

In certain clinical trials, the primary outcome has involved a target value of DAS 2.6 (3) or 1.6 (53) as a surrogate for remission or direct remission (4). These trials were not designed for regulatory approval, and did not include biologic agents. The Finnish Rheumatoid Arthritis Combination Therapy Trial (FinRACo) had a remission endpoint. In this trial, work disability rates were significantly lower in patients who received combination methotrexate, sulfasalazine, hydroxychloroquine plus prednisolone versus patients who received single DMARD therapy with or without prednisolone (54). Furthermore, if patient inflammation was controlled to a status of remission at 6 months, 4 1/2 years later no patient was receiving work disability payments. By contrast, 22% of patients who had ACR 20 or 50 responses, and 54% of patients who did not have ACR 20 responses, were receiving work disability payments (55). The TICORA study documented that a strategy of intensive tight control of RA led to a significantly better status compared to traditional therapeutic strategies in articular, functional, and radiographic outcomes over 18 months (53). These data provide strong evidence that "target control" or remission is associated with better outcomes than ACR 20 or ACR 50 responses.

In theory the goal of treatment of any disease is a "cure" or "remission." As noted, "cure" is not yet possible in diseases characterized by dysregulation of normal host control mechanisms, such as RA, hypertension, diabetes, and most other chronic non-infectious diseases, as the mechanisms of dysregulation remain poorly understood (56). Nonetheless, although the dysregulation is incurable, "tight control" of its consequences through long-term (lifetime) therapy results in lesser vascular damage in diabetes (57), increased survival in hypertension (58), and improved survival in RA (9, 10). In other diseases, control of 20% or maybe of 50% of a dysregulation appears inadequate to prevent long-term damage. For example, treatment to reduce blood pressure from, say, 200/140 to 160/112, or hemoglobin A1C from 9 to 7.2 present a 20% reduction, but is not clinically satisfactory. The observation that work disability is not seen in any patient after 5 years who was in remission after 6 months of therapy, compared to one in 5 who had ACR 20 or 50 responses, supports the concept that ACR 20 or 50 responses are not adequate.

It has been suggested that favorable values of quantitative measures known to predict mortality, such as good functional status on the HAQ or a modified version or number of involved joints (28, 39), might serve as possible "target values" for therapeutic interventions, as in the management of diabetes and rheumatoid arthritis.
hypothesis of low DAS scores (3). Such target values could serve as alternatives to remission criteria or specific changes in the ACR Core Data Set index to identify a favorable response to therapy (6).

In addition to evidence that partial control of Core Data Set and DAS measures may be associated with progression of damage in RA, there exist some intrinsic problems with these measures which may compromise their capacity to assess patients with RA. Joint count measures of swelling and tenderness may be normal in patients with histologic synovitis (59-61), and/or abnormal ultrasound and MRI scan results. Furthermore, in a study of leflunomide compared to methotrexate and placebo, measures of swollen and tender joints improved in patients who received placebo, while patient self-report measures of pain, functional disability on the HAQ or modified HAQ (MHAQ) and CRP did not improve (13). Indeed, the relative efficiencies of patient measures were greater than those for joint tenderness and swelling (62), perhaps suggesting that the investigator’s desire to influence disease activity may actually influence the measure.

Problems with other core Data Set measures are also seen. A normal ESR is seen in 40%-70% of patients (23, 63), and the ESR tends to be stable over the long-term course of RA (64). Global scores are often unchanged despite significant disease progression (65, 66). Patient questionnaires and radiographic scores appear to be the only effective Core Data Set markers to document long-term progression of RA. It appears that absolute treatment goals, rather than relatively greater suppression of inflammation at a level of 20%–50% compared to placebo or other therapies, will provide more satisfactory improvement of outcomes for most people with RA. More data are needed to analyze the levels of suppression of inflammation that will prevent or slow future damage, perhaps short of remission itself, which is difficult to achieve in all patients. Regulatory policies might be altered to require more stringent absolute therapeutic goals in formal RA clinical trials, and inclusion criteria might be modified to allow more patients to participate, thus contributing toward greater advances for the rheumatology community, and above all, patients with RA.

References

23. SOKKA T, PINCUS T: Most patients receiving routine care for rheumatoid arthritis in 2001 did not meet inclusion criteria for most recent clinical trials or American College of Rheumatology criteria for remission. J Rheumatol 2003; 30: 1136-46.
27. VAN DER HEIDJE DMFM, VAN HOF M, VAN RIEL


56. Veterans Administration Cooperative Study on Antihypertensive Agent: Effects of treatment on morbidity in hypertension: II. Results in patients with diastolic blood pressure averaging 90 through 114 mm Hg. JAMA 1970; 213: 1143-50.


