Subject Review

Rheumatoid Arthritis: Can the Long-Term Outcome Be Altered?

SHERINE E. GABRIEL, M.D., HARVINDER S. LUTHRA, M.D., Division of Rheumatology and Internal Medicine

Although several agents (for example, intramuscularly administered gold, auranofin, D-penicillamine, hydroxychloroquine, and methotrexate) are of clinical benefit in the management of rheumatoid arthritis (RA), their effect on the long-term outcome of the disease is controversial. Assessment of the influence of therapeutic interventions in RA is difficult because the natural history of the disease remains poorly defined and unpredictable, and neither the traditional clinical and laboratory measurements of inflammation nor radiographic analyses of progression of joint destruction provide an accurate estimate of the long-term outcome of RA. Furthermore, there is little evidence that second-line agents yield benefits beyond 3 years. Therefore, adequately tested comprehensive measures should be used in large, long-term, multicenter controlled clinical trials to determine whether the long-term outcome of RA can be altered.

In patients with rheumatoid arthritis (RA), there is little doubt that the so-called remission-inducing agents—that is, intramuscularly administered gold, auranofin, D-penicillamine, hydroxychloroquine, and methotrexate—are of clinical benefit. 1 A much more controversial and difficult issue, however, is whether these agents actually affect the long-term outcome of this disease.

The resolution of this controversial issue is impeded by several factors, including the natural history of RA, the inadequacy of current measures of disease outcome, and the usefulness of historical evidence for determining long-term efficacy. In this article, we discuss these factors in an attempt to answer the question, Can the long-term outcome of RA be altered? This review is intended to be selective and provocative rather than comprehensive.

NATURAL HISTORY

Because the major determinants of the onset and the course of RA are unknown, accurate assessment of the outcome of therapeutic interventions is difficult, if not impossible. Of the extensive literature on the natural history of this disease, the historical study by Short and associates is perhaps the most informative. 2-5 Between 1930 and 1936, 300 patients with arthritis were entered into this prospective study. Of these patients, seven were excluded from analysis because of rheumatic fever or infectious arthritis. The 293 remaining patients were compared with 293 age- and sex-matched control subjects. The study design had several limitations. Selection of patients was biased by many factors. All patients included in this study were inpatients. Any patient older than 12 years of age with a history of “RA” was included. Generally, a low-income population was studied. Finally, patients with inflammatory bowel disease, psoriasis, and spondylitis were included. Analysis of the selection criteria reveals that the patient group was not generalizable to patients with RA as a whole and was likely not specific enough to include only patients with RA. Selection of the control subjects was likewise biased. The first 92 subjects were obtained from outpatient departments or inpa-
tient wards of the Massachusetts General Hospital, and the other 201 control subjects were selected from various groups of hospital staff. Although the investigators excluded persons who had any evidence of joint disease in the history or on physical examination, this method of selection was certainly suboptimal. The results must therefore be interpreted in light of these serious design flaws.

The authors reviewed their results in two separate stages. The first was a retrospective analysis in which they classified patients by differing course patterns based on the duration of attacks and remissions before entry into the study. They identified 213 patients as having progressive disease and the 80 others as having intermittent disease. The group with intermittent disease was further subdivided into four subgroups (Fig. 1): patients who had brief (less than 1 year) attacks with brief remissions, patients who had brief attacks with long (more than 1 year) remissions, patients who had long attacks with brief remissions, and patients who had long attacks with long remissions. The first two groups accounted for approximately 89% of the population studied, and the last two groups constituted only 11%.

The second stage of analysis involved two prospective follow-up studies (in 1947 and 1954) of the initial patient group. The overall results are summarized in Table 1. A similar percentage of patients seemed to be in remission during each of the three evaluations; however, these numbers did not represent the same patients. This observation suggests that the frequency of long-term remissions was low and that new exacerbations and remissions were occurring. Also, the rate of improvement seemed to decrease with increasing duration of follow-up. Although 54.7% of patients had improvement in 1947 (a mean follow-up interval of 9½ years), only 35% of patients were in this category by 1954 (a mean follow-up interval of 14 years). The authors commented that even though patients who had improvement early in the course of the disease were able to maintain this status, improvement rarely followed a failure to demonstrate a favorable course for 10 years or more. Although this stage of the analysis excluded patients with spondylitis, 13 patients with inflammatory bowel disease or psoriasis were included. Several factors thought to be associated with either intermittent or progressive disease were identified (Table 2).

Another contribution to our understanding of the natural history of RA was provided by Masi and colleagues, who evaluated articular patterns in 50 patients with early RA by recording tenderness and swelling and studying these factors for a mean of 5.7 years. Three articular patterns were described (Fig. 2): (1) A monocyclic pattern, which consisted of a single cycle with complete remission for at least 1 year, occurred in approxi-

---

### Table 1.—Results of Follow-Up of 293 Patients With Rheumatoid Arthritis

<table>
<thead>
<tr>
<th>Course of disease</th>
<th>Year of follow-up* (% of patients)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1937</td>
</tr>
<tr>
<td>In remission</td>
<td>17.2</td>
</tr>
<tr>
<td>Slight to moderate improvement</td>
<td>36.9</td>
</tr>
<tr>
<td>No change</td>
<td>26.8</td>
</tr>
<tr>
<td>Progression</td>
<td>19.2</td>
</tr>
<tr>
<td>Total no. of patients</td>
<td>239</td>
</tr>
</tbody>
</table>

*Patients were entered into this prospective study between 1930 and 1936.
Modified from Short and associates. By permission of Harvard University Press.

---

Fig. 1. Subgroups of 80 patients with intermittent rheumatoid arthritis, based on duration of attacks and remissions. Numbers at bottom denote years of illness. (Data from Short and associates.)
More recently, several studies have attempted to determine the prognosis of patients with RA. In general, the conclusions from all these studies are surprisingly similar. They suggest that most patients with RA have a waxing and waning, polycyclic course characterized by remissions and exacerbations. Although the remissions may be prolonged, lasting several months or, in some cases, several years, they are more typically brief and the trend is toward continuing progression, disability, and a substantial decrease in survivorship.

Although evidence exists that some features of RA—that is, sustained high-titer rheumatoid factor, rheumatoid nodules, and extra-articular disease—connote a worse prognosis, the natural history of RA as a whole remains poorly defined. Therefore, assessment of the long-term outcome of any therapeutic measure has obvious limitations.

**MEASUREMENT OF OUTCOME**

Objective measurement of the long-term effects of therapeutic interventions in RA is extremely difficult. Estimates of improvement have traditionally been focused on “objective” measures of disease activity. The standard of such objective measurement has been radiographic progression.

**Radiographic Progression.**—The value and limitations of radiographic analysis as a measure

---

**Table 2.—Predictive Factors in a Study of Rheumatoid Arthritis**

<table>
<thead>
<tr>
<th>Factors that predicted progressive disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polyarticular small-joint involvement at onset</td>
</tr>
<tr>
<td>Short duration of arthritis (&lt;1 year)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factors that predicted intermittent disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute, severe onset</td>
</tr>
<tr>
<td>Age &lt;40 years at onset</td>
</tr>
<tr>
<td>Monoarticular involvement at onset</td>
</tr>
<tr>
<td>Fever</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factors with no predictive value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
</tr>
<tr>
<td>Presence of premonitory symptoms</td>
</tr>
</tbody>
</table>

Data from Short and associates.²⁰

---

Approximately 20% of the patients. (2) A polycyclic pattern was noted in most of the patients (approximately 70%). This pattern was further subdivided into intermittent and continuing subtypes. The intermittent group had two or more cycles separated by complete remission for at least 6 months, and the continuing group had continuous disease without complete remission or progression. (3) The progressive pattern, characterized by increasing numbers of joints involved without complete remission, was observed in approximately 10% of the patients.

The authors suggested that the articular patterns were determined early in the course of the disease and were related to certain, as yet undefined, host determinants. Certain entry variables were significantly correlated with monocyclic as opposed to progressive courses of joint involvement. These included male gender, less than 13 painful or tender joints, absence of malaise or weight loss, seronegativity, being in the “probable” class of diagnosis on the basis of the American Rheumatism Association criteria,⁶ and having normal results of a physical examination based on the New York criteria.⁸ The presence of 20 or more painful or tender joints or 10 or more upper joints that were swollen was significantly correlated with a progressive course.

The designs of these two studies were quite different. In the first study, a heterogeneous group of patients treated with conservative therapy alone were followed up for a long period; in the latter study, a well-defined group of patients with early RA were under surveillance for a short period. Although many authors have speculated on the natural history of RA, these studies provide no additional valuable prospective information on this issue.⁴⁻⁹⁻¹¹

---

Fig. 2. Articular patterns in 50 patients with rheumatoid arthritis (RA). (From Masi and associates.⁸ By permission of Technical Publishing Company.)
of outcome of RA have recently been reviewed by Sharp.22 Although the assumption among clinicians seems to be that second-line agents delay radiographic progression, few controlled studies have addressed this topic, and considerable disagreement is found among them. A recent review that critically evaluated the published evidence that drug therapy can slow radiographic progression of RA found many methodologic deficiencies.23 Inadequacies in sample size and duration of treatment, variations in drug dosages, and lack of standardization of the methods used to assess radiographic change were the major areas of concern. Overall, the evidence suggested that both gold and cyclophosphamide can retard radiographic progression, but too few technically adequate studies were available to permit even provisional conclusions about the other agents reviewed (antimalarials, D-penicillamine, and other cytotoxic agents). The studies suggested that therapy must be initiated early to influence radiographic evidence of damage.

The long-term effects of combination therapy are yet to be evaluated, but a report by McCarty and Carrera24 described "recortication" of erosions in nine patients treated with cyclophosphamide, azathioprine, and hydroxychloroquine. Another recent article attempted to define long-term progression of joint damage in RA25 by using Larsen's scores. Of 50 patients with RA who were followed up radiographically for 10 years, 48 had deterioration on the basis of the total score (mean increase, 13% of maximum). Although the authors noted an increase in joint damage among the seropositive patients, no other predictive factors could be identified from the study.

Scott and co-workers26 reported two complementary studies in which the progression of radiographic changes in RA was evaluated—one was short term (for 1 year), and the other was long term (for 10 years). Both studies involved a large group of patients (64 and 88, respectively) who received second-line agents and no placebo group. Both studies demonstrated a gradual reduction in radiographic progression with time. The authors emphasized the divergence between deterioration in radiographic features and improvement in clinical factors.

More recently, Fries and associates27 conducted a randomized controlled study in which several radiologic techniques were evaluated for the assessment of progressive joint destruction in RA. Several important conclusions were drawn from this study. Erosions and joint space narrowing were shown to contribute independent information. Although the number of joints evaluated made little difference, the number and the expertise of the observers were critical, as was careful attention to the selection of a scoring method and to the measurement of study endpoints.

In another review, Pullar and colleagues28 reported their results in 67 patients with RA (57 who received various second-line agents and 10 control subjects who refused second-line therapy) who were followed up clinically and radiographically for 2 years. They demonstrated radiographic deterioration in all patients, but a trend toward slowing of the rate of development of erosions was noted in the patients who received gold and D-penicillamine in comparison with control subjects.

Whether erosion count is a valid measurement of outcome in RA trials was recently questioned.29 In that study, roentgenograms of the hands of 29 patients with RA taken at various times during their disease course, from diagnosis to 4 years later, were interpreted by two expert radiologists. Interobserver and intraobserver agreement was poor when the roentgenograms were studied in random order and even when studied with knowledge of the chronologic sequence (agreement, 40 to 50%). Agreement was increased to 70 to 80% if the presence or absence of erosions was known. The authors concluded that erosion count was not a valid indicator of outcome, even when senior, experienced bone radiologists were the observers.

Sharp and co-workers30 recently reported the results of a large collaborative study of the reproducibility of radiographic analyses of hand and wrist abnormalities in patients with RA. With use of different methods, 13 experienced observers evaluated 41 coded roentgenograms, which disclosed a wide variety of abnormalities. The study showed good general agreement in the scoring of roentgenographic abnormalities; correlation coefficients were greater than 0.850 for approximately two of three comparisons.

In summary, radiographic progression, when assessed by experienced bone radiologists using a well-defined reproducible method of analysis, does provide an indication of progressive destructive RA. Although such an indicator is valuable, its correlation with long-term outcome is impre-
exercise and the changes seen are often irreversible. Ideally, a measurement of outcome in RA should more closely estimate the outcome of disease during both the early reversible stages and the later stages. Such a measurement should provide a valuable guideline for therapy as well as be a useful indicator of the long-term influence of the disease.

Other Studies.—Other estimates of improvement used repeatedly in clinical trials of RA include 50-foot walk time, grip strength, morning stiffness, joint count, and laboratory values (erythrocyte sedimentation rate, hemoglobin, rheumatoid factor titer, and C-reactive protein). It is becoming increasingly clear that such data are often conceptually and empirically limited in terms of their ability to measure improvements in health. Several authors have addressed the diurnal, sequential, and cyclic variations with use of these clinical measures as well as the significant intraobserver and interobserver variability and their poor reproducibility. Other methods of monitoring disease activity such as thermography and multivariate analytical techniques that consider several clinical and laboratory factors have not proved to be widely useful. Many acute-phase reactants have been proposed as markers of disease activity in RA, such as erythrocyte sedimentation rate, hemoglobin, albumin, C-reactive protein, serum amino-terminal type III procollagen peptide, serum amyloid A protein, urinary neopterin, and others. Recently, several studies have suggested that C-reactive protein, determined by quantitative assays, is a valuable marker for disease activity in RA and particularly for ankylosing spondylitis and Reiter’s disease. Furthermore, it may have better sensitivity and specificity than the erythrocyte sedimentation rate.

The value of clinical estimates of improvement was addressed by Koran in an extensive review that evaluated the reliability of many signs and procedures and the diagnostic and therapeutic judgment among physicians. Major disagreements in diagnostic approaches were found among physicians. In another study, Kirwan and associates evaluated the clinical judgment of 48 randomly selected British rheumatologists concerning the progress of 50 simulated case histories. The rheumatologists differed appreciably in their assessment of the progress of the disease, even when questions with “yes” or “no” answers in regard to clinically important changes were considered.

The reliability and validity of the aforementioned clinical and laboratory measures are surprisingly low. Furthermore, such factors tend to measure the disease process rather than the outcome. Outcome has been defined as the “end result,” whereas the disease process is simply “what happens along the way.” For example, disability is an outcome, and erythrocyte sedimentation rate measures a part of the process. Because the ultimate goal of medical care is to improve the overall outcome of patients, measures of outcome clearly would be of more value than measures of disease process. In fact, measures of disease process are useful only to the extent that they may or may not approximate the outcome. The results of the same therapeutic intervention can have different interpretations, depending on whether measures of outcome or of process are used (Table 3). Therefore, the accuracy of attempts to measure long-term outcome is limited by the inadequacy of the measuring tools.

Recently, experience with more comprehensive endpoint measurements that more completely define health status has been reported. Several health status “instruments” were recently developed for use in rheumatology. These are conceptually relevant, and extensive methodologic study has shown that their reliability and validity are, in fact, equal to or greater than those

<table>
<thead>
<tr>
<th>Process</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>In rheumatoid arthritis, 20 mg of prednisone daily is clinically and statistically superior to placebo for reducing synovitis, decreasing morning stiffness, increasing walking speed, decreasing the number of tender joints, improving grip strength, and decreasing the sedimentation rate</td>
<td>In rheumatoid arthritis, 20 mg of prednisone daily, as compared with nonsteroid treatment, leads to increased mortality, an increased rate of development of disability, increased symptoms due to long-term side effects, more hospitalization, and increased direct and indirect costs of disease</td>
</tr>
</tbody>
</table>

From Fries. By permission of the American Rheumatism Association Educational Foundation.
of many of the traditional methods. They also fulfill several criteria necessary for measurement instruments: they permit quantification, have established validity and reliability, and demonstrate both standardization of data collection and measurement precision.

The question of whether these instruments are sensitive enough to measure the differences in outcome that might be expected to occur in a specific trial was answered in a recent article by Meenan and colleagues, who described the use of a self-administered health status questionnaire in a randomized double-blind 21-week comparison of placebo, orally administered gold, and injectable gold in patients with RA. The outcome was assessed by standard clinical measures and by the Arthritis Impact Measurement Scales (AIMS). Data from the clinical and health status measures produced highly similar conclusions—that is, parenterally administered gold is more efficacious than orally administered gold, and orally administered gold is more efficacious than placebo. The results strongly suggested that this health status measure was able to detect meaningful improvements in outcome in the clinical trial setting. Discriminant analysis showed that information from four clinical factors (swelling, tenderness, grip strength, and overall patient assessment) enabled investigators to assign subjects to their correct drug category in 42% of cases, and information from the AIMS did so in 43%. When the two sets of data were considered collectively, assignment of subjects was correct in 71% of cases on the basis of information from clinical factors, in 79% with information from the AIMS, and in 79% if both were used together. This last finding is of interest. Treatment discrimination by these two types of measures was additive; thus, they might be used effectively in tandem to assess the efficacy of a specific intervention.

Thus, neither the traditional clinical and laboratory measurements of inflammation nor radiologic analyses are able to provide an accurate estimate of the long-term outcome of RA or guide therapy adequately in RA. Although the new methods of functional assessment attempt to approach this goal more closely, further study and experience are needed before their efficacy can be established.

CURRENT EVIDENCE FOR LONG-TERM EFFICACY OF SECOND-LINE AGENTS AND CORTICOSTEROIDS

Although the treatment of RA involves a comprehensive program of physical therapy, education, nonsteroidal anti-inflammatory agents or aspirin, and social support, these interventions will not be discussed as they are not generally thought to alter the course of the disease and likely do not substantially affect the outcome.

Controlled clinical trials have proved the efficacy of several second-line agents for the treatment of RA. In general, however, these studies all have the same weaknesses—that is, the duration of the treatment periods and the extent of the follow-up periods are inadequate to assess long-term benefit accurately.

The earliest placebo-controlled trials of second-line agents were those by Fraser in 1945 and the Empire Rheumatism Council in 1960 and 1961, in which gold therapy for RA was evaluated. Consequently, the beneficial effect of gold has been irrefutably established. Current evidence suggests that the benefit from intramuscularly administered gold compounds may continue for up to 3 years, as determined by delayed radiographic progression. The duration of response beyond 3 years remains poorly defined; in fact, some studies have shown loss of benefit after 4 to 6 years of therapy.

Similarly, auranofin, a relatively new orally administered gold compound, has been demonstrated to be beneficial in RA. Many studies provide evidence for continued benefit at 24 months. Although some reports suggest a continued response from auranofin for up to 5 years, only a few such long-term studies have been reported. Therefore, the exact duration of response beyond 2 years remains undefined.

Hydroxychloroquine is perhaps the best tolerated second-line agent, being limited primarily by associated retinal toxicity. Although clinical trials with this medication have also been generally less than 2 years in duration, a recent study by Runge suggested a more lasting pattern of response to this agent. Using life-table retrospective analysis, he showed that relapses occurred at a continuous rate during 50 months of gold therapy, whereas the rate of relapse plateaued after 20 months of treatment with hydroxychloroquine. In addition, a second analysis revealed that the rate of discontinuation of treatment was
lower with hydroxychloroquine (25% at 1 year of treatment) than with gold and D-penicillamine (45% at 1 year of treatment). 81

A 2-year prospective double-blind study 82 compared hydroxychloroquine and D-penicillamine singly and in combination. Although D-penicillamine seemed superior to hydroxychloroquine at the 6-month and 12-month evaluations, thereafter the efficacy of D-penicillamine decreased linearly such that at 24 months only 11% of the patients were “markedly improved” in comparison with baseline. In contrast, 50% of the patients who received hydroxychloroquine remained improved after 2 years, and 25% were “markedly improved.” This study also revealed that combination therapy with these two agents was actually less efficacious than therapy with either drug used singly. No reduction in radiographic progression of RA was demonstrated in any of the three treatment groups.

Although D-penicillamine has been shown to be an effective antirheumatic agent, 83-86 its efficacy being approximately equivalent to the other second-line agents, 87 patient compliance with long-term treatment has generally been poor because of the development of toxicity. 88,89 In addition, a recent 5-year prospective comparison of D-penicillamine with other antirheumatic agents demonstrated a failure of D-penicillamine to arrest radiographic progression of RA. 90

Several cytotoxic or immunosuppressive agents have proved to be efficacious for the treatment of RA. 91 Cyclophosphamide has been shown to delay radiographic progression of the disease, 24,92,93 but no benefit lasting longer than 18 months has been established. Because of their potential toxicity, these compounds have generally been reserved for patients in whom therapy with two or more sequential agents has failed.

Recently, Wolfe and Hawley 94 evaluated the frequency and duration of “remission,” as defined by the American Rheumatism Association remission criteria, 95 in 458 patients with RA who were followed up for 1,131 patient-years. They found that 18.6% of gold treatment courses and 16.7% of D-penicillamine treatment courses resulted in remission and that 13.6% of patients who received no therapy experienced remission. The median duration of remission was 10 months, and only 7.8% of the observed time course was spent in remission. In patients receiving therapy, female sex, disease onset before age 60 years, and early development of erosions were associated with a decreased likelihood of remission.

The role of glucocorticoids in the treatment of RA remains undefined. Since the original report of the use of compound E 96 and the realization of the side effects, 97 few clinical trials have examined the usefulness of systemic glucocorticoid therapy in RA. The earliest of these studies indicated that 10 mg or more of prednisolone daily resulted in a greater and more rapid effect on RA than did 3.9 g of aspirin daily. 98 The authors demonstrated that the prednisolone-treated patients had significantly fewer new erosions of the hands and feet than those in the aspirin-treated group. A follow-up report that analyzed new erosions in both patient groups during a period of 4 years or longer showed that prednisolone treatment decreased both the proportion of patients with RA in whom new erosions developed and the rate of development of new erosions. 99 Their data, recently reviewed by Masi, 100 showed that new erosions developed in 94% of the 34 aspirin-treated patients (at a rate of 3.25 new erosions/patient per year) but in only 51% of the 39 prednisolone-treated patients (at a rate of 0.77 new erosion/patient per year). As noted by Masi, however, no follow-up data regarding functional status or side-effects were reported.

A more recent study 101 evaluated low-dose daily prednisone therapy (5 mg) versus placebo when used with other drugs in 18 patients with RA. After 24 weeks, all patients were treated with placebo for 8 additional weeks. Although no significant difference in response was demonstrated between the prednisone and placebo groups at 24 weeks, significant deterioration occurred when the prednisone therapy was abruptly discontinued.

The generally accepted uses of glucocorticosteroids are to manage the serious extra-articular manifestations of RA and the serious adverse effects of the second-line agents (for example, gold-induced colitis and nephritis, D-penicillamine-related thrombocytopenia, and lupus-like reactions) and to provide a potent anti-inflammatory “bridge” during the lag phase of weeks to months before the onset of the effect of the second-line agents. The questions of whether glucocorticoids can independently act as disease-modifying agents and whether daily therapy with low-dose prednisone can positively influence the long-term outcome in RA remain unanswered.
In summary, of the drugs that have been generally accepted as disease-modifying agents, only intramuscularly administered gold and cyclophosphamide have been demonstrated to retard the radiographic progression of joint destruction, and little evidence shows that any of the antirheumatic agents provide benefit beyond 3 years.

**DISCUSSION**

On the basis of the preceding review, several questions can be posed.

1. If we are indeed unable to alter the long-term outcome of RA, is it worthwhile to treat patients with disease-modifying drugs?

Even if no long-term benefit can be demonstrated from these agents, improvement in the quality of life and disease suppression, even if temporary, are of value in the management of RA. This goal, of course, must be weighed against the toxicity of the agents used to achieve it.

2. What should be the goal of second-line drug therapy in RA?

The design and analysis of clinical drug trials would be improved if a more accurate definition of the goal of drug therapy in RA was established at the outset. Second-line drug therapy in RA should be aimed at altering disease outcome. The alternative “outcome goals” can be defined as follows:

   a. Cure (eradication or reversal of all disease)
   b. Control (retarding [less than 100% efficacy, partial remission] or arresting [100% efficacy, complete remission] disease progression)
   c. Relief (symptomatic improvement without affecting the disease process)

3. What is the most meaningful way of measuring the outcome of therapeutic interventions in RA?

Comprehensive measures that incorporate adequately tested functional instruments with clinical, laboratory, and radiographic data are necessary to evaluate outcome optimally. In addition, the search must be continued for more accurate measures of outcome.

4. How can we design clinical trials to help answer the question about the potential for altering the long-term outcome of RA?

The controlled clinical trial is the final step in determining the efficacy, safety, and usefulness of drugs for the rheumatic diseases. It is the only way to answer the question, Can the long-term outcome of RA be altered? Trials must be randomized and prospective and must incorporate both short- and long-term goals. Because the expected benefits in RA are often small, achieving adequate power necessitates the enrollment of large numbers of patients in studies. For example, when evaluating two treatments, one that is 40% efficacious and one 60% efficacious, one would need approximately 150 patients in each group to achieve a power of 0.9 (Fig. 3).

Performing this study with the proper patient population at one medical center may take many years and would be logistically and financially impractical. Therefore, multicenter trials involving several medical centers would be necessary to obtain adequate power.}

![Fig. 3. Numbers of patients needed in each treatment group to complete a clinical trial that would allow a 90% chance of demonstrating superiority of one treatment over another. For example, if treatment A induces a 40% response and treatment B a 60% response, 150 patients would be needed in each group if a trial is to have a 90% chance of demonstrating this difference. (Data from Fleiss JL. Statistical Methods for Rates and Proportions. New York, John Wiley & Sons, 1981.)](image-url)
difficult. A well-organized, cohesive group of cooperating centers, however, would be much more capable of rapidly and efficiently conducting such a study. Of course, the potential benefits of the results of such a study need to be carefully weighed against the estimated costs of conducting a study of such magnitude. Perhaps if these ideas are considered in the design and conduct of clinical trials, the answer to this most fundamental question may become apparent.

REFERENCES


57. Meenan RF: The AIMS approach to health status measurement: conceptual background and measurement properties. J Rheumatol 9:785-788, 1982
68 LONG-TERM OUTCOME OF RHEUMATOID ARTHRITIS


96. Hench PS: The reversibility of certain rheumatic and nonrheumatic conditions by the use of cortisone or of the pituitary adrenocorticotropic hormone. Ann Intern Med 36:1-36, 1952


