Limitations of Traditional Randomized Controlled Clinical Trials in Rheumatology

Theodore Pincus

Introduction

The randomized controlled clinical trial is appropriately regarded as the most rigorous method to document the efficacy of a therapy compared to another therapy or a placebo. A clinical trial allows isolation of a single variable, the test therapy, mimicking a laboratory "scientific experiment" [1]. This approach conforms to a "biomedical model" [2], the dominant paradigm of contemporary medicine. In recent years, the randomized controlled clinical trial often has been regarded in the medical literature as the *only* approach to assess the value of a new therapy according to "evidence-based medicine" [3]. However, randomized trials have many limitations, some of which are summarized in this chapter.

The earliest randomized controlled trials were conducted in the 1940s in infectious diseases such as tuberculosis [4, 5]. Clinical trials in infectious diseases have advantages over those in many other diseases, particularly chronic diseases, for several reasons. First, the target of the medication involves simple unicellular pathogens such as bacteria or fungi, rather than complex mammalian cells. Therefore, any efficacious antibiotic medication without an adverse effect is likely to benefit *all* individuals infected by the pathogen that is the target of the medication. By contrast, much greater variation is seen in responses of individuals to medications which affect mammalian cells, as seen in chronic rheumatic diseases. Second, results of a therapy in an infectious disease generally are apparent over days, weeks, or sometimes months, in contrast to years and even decades in chronic rheumatic diseases. For example, superior efficacy of penicillin versus placebo for a streptococcal sore throat can be documented definitively after 10–14 days of treatment in

T. Pincus, MD (✉)
Division of Rheumatology, Rush University Medical Center,
1611 West Harrison Street, Chicago, IL 60612, USA
e-mail: tedpincus@gmail.com

H. Yazici et al. (eds.), *Understanding Evidence-Based Rheumatology*,
DOI 10.1007/978-3-319-08374-2_8

all infected individuals, while treatment effects in a chronic rheumatic disease vary among individuals and even may indicate efficacy or no differences from placebo in groups after 6–24 months, but different outcomes after 5–10 years, as discussed in detail below.

Rheumatic diseases, as well as most noninfectious diseases, do not involve "foreign" cells as in infectious disease or chemicals that require eradication to restore homeostasis. Rheumatic diseases involve a dysregulation of normal cells and/or chemicals which may be over- or underproduced due to faulty internal signals. Similar pathogenetic mechanisms based on dysregulations are seen in many common chronic diseases such as hypertension, hyperlipidemia, or diabetes. The natural history of an untreated dysregulation is organ damage to blood vessels, kidneys or joints, or other organs.

Infectious diseases are "curable" through eradication of a foreign pathogen. By contrast, dysregulatory diseases are *incurable*, based on current knowledge. However, *control* of the dysregulation retards or prevents organ damage and indirectly prevents or reduces premature mortality associated with these diseases [6–12]. Nonetheless, long-term indefinite ongoing medication generally is required, since no therapy to eradicate the etiology of the dysregulation is available at this time.

While rheumatic diseases are similar to hypertension, hyperlipidemia, or diabetes in a pathogenesis involving dysregulation of normal components leading to organ damage [13], rheumatic diseases differ from the other diseases in several features. One important difference is that rheumatic diseases are not characterized by a single "gold standard" biomarker such as blood pressure, hemoglobin A1c, bone density, etc., that can be applied to diagnosis, assessment, prognosis, and monitoring of *all individual* patients [14]. Therefore, an index of multiple measures is needed to assess and estimate changes in the clinical status of patients with rheumatic diseases.

The discovery in the 1940s of rheumatoid factor [15, 16] in rheumatoid arthritis (RA), antinuclear antibodies (ANA) [17] in systemic lupus erythematosus (SLE), and other biomarkers led to hopes that laboratory tests could be used effectively for diagnosis and management of all individual patients with RA, SLE, and other rheumatic diseases, similar to other diseases in a traditional "biomedical model." However, more than one-third of patients with RA have a negative test for rheumatoid factor, or anti-cyclic citrullinated peptide (anti-CCP) antibodies (ACPA) [18–21], and more than 40 % have a normal erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP) at presentation [21]. More than one-third of patients with SLE have no detectable anti-DNA antibodies, anti-Smith (anti-Sm), and anti-ribonucleoprotein (anti-RNP), while a positive ANA test is found in least 10 % of the normal population [22–24].

In the absence of a single gold standard measure, as noted, a pooled index [25] is applied to most rheumatic diseases. Formal indices have been developed for RA [26–31], SLE [32–39], vasculitis [40–45], psoriatic arthritis [46–48], ankylosing spondylitis [49–53], and other rheumatic diseases. These indices generally include three types of measures from patient self-report, physical examination, and laboratory tests; data may be included in some indices, particularly in longer studies. The formal indices are used in clinical trials and other clinical research, but not widely in routine clinical care [54, 55].

Inclusion of patient history information and specific physical examination findings, e.g., joint counts, reflects that a patient history and physical examination are more significant in clinical decisions in rheumatic diseases than in many other types of chronic diseases [56]. Information from a patient history may be captured as standardized, "scientific" quantitative data, according to validated self-report questionnaires [57]. Patient questionnaires may be used effectively to guide management, document change in status, assess outcomes, and improve the quality of care in rheumatic diseases, analogous to laboratory tests in other diseases [58]. Inclusion of a specific patient questionnaire at every visit of every patient ensures that some quantitative data are collected at each encounter with minimum effort on the part of the doctor and staff [59].

A contemporary view of "evidence-based medicine" recognizes limitations of clinical trials, as presented in the chapter "Evidence-based medicine in rheumatology: how does it differ from other diseases?" and described in a number of thoughtful reports by several observers [1, 3, 60–81], as well as in some of the author's own commentaries [82–88]. A recent report from the Oxford Centre for Evidence-Based Medicine [60] noted that "While they are simple and easy to use, early hierarchies that placed randomized trials categorically above observational studies were criticized [3] for being simplistic [61]. In some instances, observational studies give us the 'best' evidence [3]. For example, there is a growing recognition that observational studies – even case-series [62] *and anecdotes* [63] can sometimes provide definitive evidence."

Recognition of limitations of clinical trials in no way denies their value as the optimal method to distinguish short- and medium-term treatment effects of a medication from another medication or placebo. Indeed, it might be optimal if most patients with a chronic rheumatic disease would have an opportunity to participate in a randomized controlled clinical trial, because of the largely "experimental" nature of most available treatments. Since a "best" therapy for an individual patient usually is not identified, a health professional must "guess" at the best treatment for most patients. In this situation, the most ethical approach might appear to randomize the patient to one of several treatments, so the individual patient has a chance to experience the "best" treatment for herself/himself [89].

Therefore, it is recognized that the methodology of the randomized controlled clinical trial often provides a framework of an optimal method to evaluate the efficacy of a therapy. However, it also is important to recognize limitations of randomized controlled clinical trials, just as there are limitations to any method to acquire knowledge in medicine or any field. The author's recognition of limitations of clinical trials is based in large part on experience in conducting more than 35 randomized clinical trials.

Limitations of clinical trials are grouped into two categories. Those resulting from issues in practical implementation in modern clinical research are termed *pragmatic* limitations. Other limitations would exist even if all pragmatic limitations could be overcome, but are weaknesses of the methodology (as exist for any methodology, as noted, but often overlooked for clinical trials) and are termed *intrinsic* limitations (Table 1).

Table 1 Pragmatic and intrinsic limitations of clinical trials in chronic rheumatic diseases

Pragmatic limitations of clinical trials	
1	A relatively short time frame in chronic diseases – sometimes too short to identify important clinical benefits or to recognize loss of efficacy
2	Inclusion and exclusion criteria may restrict eligibility to fewer than 10 % of patients with a particular diagnosis who may be considered eligible for a clinical trial
3	Differences between a medication and a placebo are required to be statistically significant but not necessarily robust – statistical significance may indicate only marginal clinical benefit
4	Clinically important differences may not be statistically significant due to insufficient numbers of patients for statistical power
5	Important variables affecting outcomes other than whether a patient was randomized to a medication versus another medication or placebo may be seen – but generally ignored in clinical trial reports
6	Traditional clinical trials with parallel designs have inflexible dosage schedules and restrict concomitant medications, although a flexible dosage schedule toward a target with multiple medications may provide optimal results
7	Surrogate markers and indices used in clinical trials may be suboptimal measures to detect changes in clinical status or predict important clinical outcomes
8	Rare side effects cannot be identified in most trials
Intrinsic limitations of clinical trials	
1	The design can greatly influence results – availability of a control group does not eliminate bias
2	Data are reported in groups – ignore possible substantial variation in groups
3	No absolute criteria for the balance of risk and benefit for the therapy – different individuals may interpret very differently and all be "correct"
4	Loss of a placebo effect in a clinical trial (although gain of more extensive care which may offset and even surpass the usual "placebo effect")

Eight Pragmatic Limitations of Randomized Clinical Trials in Chronic Diseases

Eight types of pragmatic limitations in chronic diseases are summarized below:

1. *The relatively short time frame of clinical trials in chronic diseases.*
   A prominent limitation of clinical trials in chronic diseases involves too short a time frame of observation to recognize meaningful clinical trends that develop only over longer periods. For example, a randomized controlled clinical trial in RA was conducted over 48 weeks to compare results of 3 regimens – methotrexate monotherapy, auranofin (oral gold) monotherapy, and a combination of methotrexate and auranofin [90]. No significant differences were found between results with any of these three regimens (Fig. 1) [90].

   A similar conclusion was reported from a far more extensive meta-analysis of 66 clinical trials reported in 1990 concerning the efficacy of disease-modifying antirheumatic drugs (DMARDs) in the treatment of RA [91] (Fig. 2). This meta-analysis included 117 treatment groups: 11 for antimalarial drugs

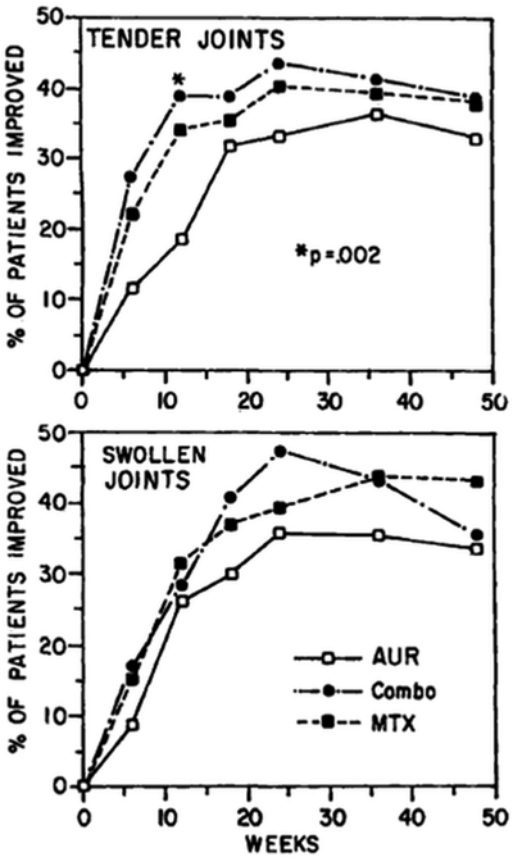

Fig. 1 Results of a randomized clinical trial in 297 patients with rheumatoid arthritis treated with auranofin (*AUR*), methotrexate (*MTX*), or auranofin plus methotrexate (*Combo*) [90]. The figure illustrates percentages of patients with ≥50 % meaningful improvement in tender or swollen joints. Final results showed no significant differences between the three groups over one year

(e.g., hydroxychloroquine), 23 for auranofin, 29 for in efficacy injectable gold, 7 for methotrexate, 19 for d-penicillamine, 6 for sulfasalazine, and 22 for placebo. The meta-analysis indicated no significant differences in efficacy between sulfasalazine, d-penicillamine, methotrexate, and injectable gold (Fig. 2) [91], i.e., that the efficacy of methotrexate for RA was equivalent to hydroxychloroquine, sulfasalazine, d-penicillamine, and injectable gold.

Results of the meta-analysis did not appear translated into actual clinical practice over 5 years in an observational study of duration of treatment courses of DMARDs in 7 rheumatology practices reported in 1992 [92] (Fig. 3, Panel a). Duration of treatment courses in an incurable chronic disease such as RA can serve as a composite measure of effectiveness and safety of a medication. A formal analysis of estimated duration of continuation of 1,083 courses of 6 DMARDs over 60 months in 477 patients with RA indicated that approximately 80 % of methotrexate courses were continued after 2 years, compared to 50 % of courses of hydroxychloroquine, penicillamine, parenteral gold, and azathioprine and only 20 % of courses of oral gold (Fig. 3, Panel a). After 5 years, approximately

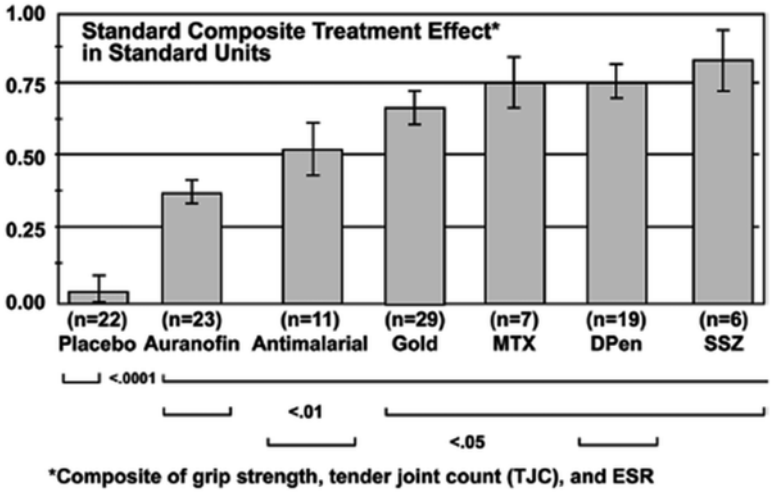

Fig. 2 Standard composite treatment effect (in standard units). Meta-analysis of 66 clinical trials reported in 1990 concerning the efficacy of DMARDs in the treatment of RA [91]. This meta-analysis included 117 treatment groups: 11 for antimalarial drugs (e.g., hydroxychloroquine), 23 for auranofin, 29 for injectable gold, 7 for methotrexate, 19 for d-penicillamine, 6 for sulfasalazine, and 22 for placebo. All drugs have greater efficacy than placebo in the management of RA, determined according to a composite of grip strength (a measure of effectiveness of grip), tender joint count, and erythrocyte sedimentation rate, adjusted for disease duration, trial length, initial tender joint count, and blinding. In these analyses, no significant differences were seen between sulfasalazine, d-penicillamine, methotrexate, and injectable gold (From Felson et al. [91] with permission)

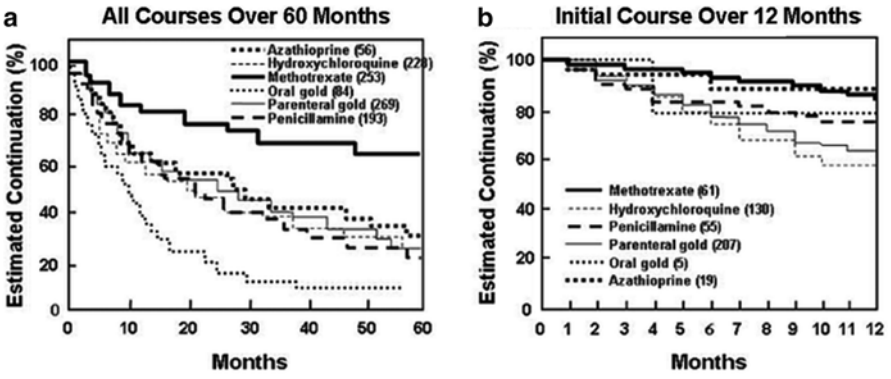

Fig. 3 (**a**) Estimated continuation of all 1,083 courses of DMARDs in 532 patients with rheumatoid arthritis over *60 months*. Differences between methotrexate and all other drugs, as well as between oral gold (auranofin) and all other drugs, are statistically significant ($P<0.001$), while differences among other drugs are not significant. (**b**) Estimated continuation of 477 courses of the *initial* DMARD used in the same 532 patients over *12 months*. Differences between methotrexate versus oral gold (auranofin) are not statistically significant and are considerably less apparent than in A, in which estimated continuation was studied for *all* courses over *60 months* [92]

60 % of the methotrexate courses were continued versus approximately 20 % of the hydroxychloroquine, penicillamine, parenteral gold, and azathioprine courses, and virtually no course of oral gold (Fig. 3, Panel a) [92].

The data from the observational study were analyzed over only *1 year* for the *initial* 447 DMARD courses, conditions that mimic clinical trials (Fig. 3, Panel b), in contrast to the above analyses of *all* DMARD courses over *5 years* (Fig. 3, Panel a) [92]. Continuation rates of courses of all 6 DMARDs were similar, including no difference between methotrexate versus parenteral versus oral gold (auranofin) (Fig. 3, Panel b), as seen in the clinical trial (Fig. 1).

The absence of statistically significant differences between DMARD courses over 1 year (seen in Fig. 3b) mimics results of clinical trials in Figs. 1 and 2 but differs considerably from the results seen in actual clinical care over 5 years (Fig. 3a). Therefore, results of both the clinical trials and observational study are accurate and "correct." However, the accurate data in the clinical trials and meta-analysis were not translated into long-term clinical care over 5 years, and the clinical trial results were *not* applicable to routine clinical care.

These observations suggest caution in interpretation of data from clinical trials to physicians for routine care. Nonetheless, in 2008 (16 years after publication of the report of differences between results of clinical trials and clinical care [90]), a "systematic review" of DMARDs in the principal journal for internists, *Annals of Internal Medicine*, concluded that there was "moderate evidence that sulfasalazine, leflunomide, and methotrexate were equivalent in efficacy, with no obvious major differences in adverse events and discontinuation rates among these three DMARDs" [93].

This conclusion differs from contemporaneous clinical care in the international QUEST-RA database of many countries (Table 2), in which methotrexate was taken by 83 % of patients, sulfasalazine by 43 %, leflunomide by 21 %, and biological agents by 23 % [94]. These patterns were seen in countries in which patients do not pay for medications [94], so they could be explained only in small part on the basis of costs. A strict methodologist may conclude that the clinicians were in error and not practicing "evidence-based medicine," since the systematic review concluded that the three agents were similar in efficacy and adverse events. However, if the conclusion of the systematic review were accurate, comparable usage of the 3 DMARDs might be expected in routine care, but that is not seen. These findings again indicate that data from short-term clinical trials may provide less accurate information about long-term results of therapies than long-term observational studies, as a result of limitations of the clinical trial methodology [87].

Limitations of a short time frame also are seen in a trial conducted in patients with polymyositis to compare therapeutic efficacy of a combination of prednisone plus azathioprine versus prednisone monotherapy (plus placebo) [95, 96]. The initial report concerning this clinical trial indicated no differences between the two groups after 3 months of treatment, according to three measures, i.e., days to normalize the creatinine phosphokinase (CPK) muscle enzyme, change in the muscle strength score, and reduction of inflammation on the muscle biopsy

Table 2 The use of disease-modifying antirheumatic drugs (DMARDs) in the QUEST-RA countries

Country	Delay to start DMARDS, months, median	DMARD exposure years, mean	Selected DMARDs ever taken; percentage of patients in the QUEST-RA study per country					
			Prednisone (%)	MTX (%)	HCQ (%)	SSZ (%)	LEF (%)	Any biological agent (%)
Argentina	13	3.7	83	68	49	***6***	16	3
Denmark	10	7.9	43	85	39	64	11	23
Finland	7	14.4	74	85	**74**	**84**	21	17
France	8	9.9	83	86	55	49	**42**	**53**
Germany	15	8.4	54	78	30	36	25	29
Ireland	11	6.3	71	**92**	***15***	33	24	41
Italy	9	7.1	69	79	42	14	31	26
The Netherlands	5	8.1	***26***	91	28	35	6	19
Poland	4	7.2	69	87	34	60	18	8
Serbia	11	6.6	**88**	69	55	17	7	***2***
Spain	14	7.3	67	82	43	29	34	27
Sweden	12	8.8	66	83	34	62	9	31
Turkey	12	8.9	69	88	27	61	22	7
UK	12	7.9	51	***67***	39	46	***4***	16
USA	9	7.9	77	85	49	12	19	33
Total	9	8.1	66	83	41	43	21	23

Adapted from Sokka et al. [94]
The highest percentage for each drug is indicated in bold and the lowest in bold italics
DMARD disease-modifying antirheumatic drug, *HCQ* hydroxychloroquine, *LEF* leflunomide, *MTX* methotrexate, *QUEST-RA* Quantitative Standard Monitoring of Patients with Rheumatoid Arthritis, *SSZ* sulfasalazine

Table 3 Comparison of results in treatment of polymyositis with prednisone + azathioprine versus prednisone + placebo over 3 years according to functional grade disability [96]

Treatment	Functional grade disability		
	Onset	1 Year*	3 Years*
Prednisone + azathioprine	4.5	3.0	2.1
Prednisone only	4.1	3.6	3.0

Adapted from: Bunch [96]
*$p < 0.01$

[95]. The authors concluded that "in a controlled, prospective, randomized, double-blind study…azathioprine does not afford any therapeutic advantage when used in addition to accepted prednisone dosages in the initial management of polymyositis" [95].

The randomized controlled trial was then continued further to 3 years. After 3 years (Table 3), improvement according to functional grade disability was significantly greater for patients treated with the combination of prednisone plus

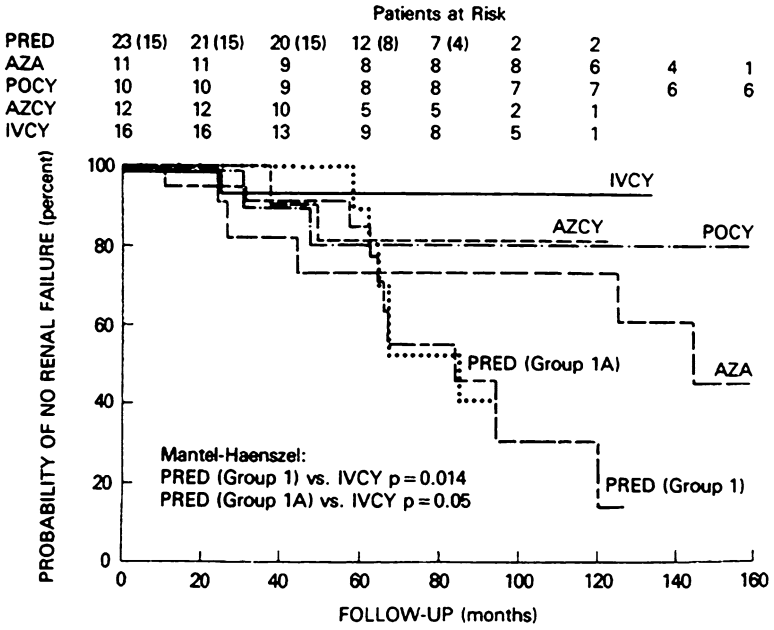

Fig. 4 Probability of maintaining life-supporting renal function in long-term randomized clinical trials of 72 high-risk patients with active SLE nephritis, according to treatment group: *PRED* prednisone, *AZA* azathioprine, *POCY* oral cyclophosphamide, *AZCY* combined oral azathioprine and cyclophosphamide, *IVCY* intravenous cyclophosphamide [97]

azathioprine versus those treated with prednisone monotherapy. In retrospect, differences were seen after 1 year according to functional status, but not according to CPK, muscle strength, or muscle biopsy. The authors concluded that "longer follow-up (3 years) has shown that the group given prednisone plus azathioprine has improved more with respect to functional disability; this group also requires less prednisone for disease control" [96].

These observations illustrate two important principles regarding analyses of treatments in a rheumatic disease: (*a*) Recognition of the possible advantages of combination second-line therapy may require periods of years, rather than months. (*b*) Differences in results of two treatments may be apparent according to measures of functional disability, rather than laboratory or biopsy data, as discussed below (see point 7 concerning surrogate markers). These principles may be relevant to studies of all rheumatic diseases.

A striking example of the importance of a long time frame in a clinical trial to assess treatment of a chronic disease is seen in a trial designed to prevent renal failure in patients with SLE nephritis using several treatment regimens, including prednisone monotherapy versus combinations of prednisone with azathioprine and/or cyclophosphamide (Fig. 4) [97]. Substantial advantages to cyclophosphamide plus prednisone were seen over 10 years, with preservation of renal function in about 90 % of patients versus only about 30 % of patients

treated with prednisone monotherapy (Fig. 4). These results established cyclophosphamide as the standard of care for SLE nephritis for at least two decades in the 1980s and 1990s. It is not widely recognized, however, that even after 4 years, renal function was preserved in more than 90 % of patients in all groups, i.e., prednisone monotherapy appeared as effective as the combination with cyclophosphamide (Fig. 4).

As a result of this clinical trial, combination therapy with cyclophosphamide plus prednisone became the standard of cate for SLE nephritis over the next two decades. However, if this trial had been conducted over only a 3-year period or less, *as is the case in more than 98 % of randomized controlled trials in rheumatology*, it would have been concluded that no advantage is seen to the combination of prednisone plus cyclophosphamide over prednisone monotherapy! Only a clinical trial conducted in relatively asymptomatic individuals with support from the intramural program of the United States National Institutes of Health (NIH) allowed 10 years of observation, which would not be possible at most clinical settings to establish a new standard of care.

2. *Inclusion and exclusion criteria may restrict eligibility to fewer than 10 % of patients with a particular diagnosis who may be considered eligible for a clinical trial.*

In theory, all individuals with a particular diagnosis should be eligible to participate in a clinical trial. This goal is more likely to be met in a short-term trial of an antibiotic to eradicate an infectious agent, rather than for a longer-term (but usually not long enough) trial of a medication that affects primarily mammalian cells. In practice, however, all clinical trials have inclusion and exclusion criteria, designed to ensure that a relatively homogeneous group of patients with sufficient disease activity and without severe confounding comorbidities are studied to document improvement.

In many, if not most, instances, inclusion criteria are rather stringent, so that only a small fraction of patients are eligible for the trial. For example, inclusion criteria in the Anti-TNF therapy in RA with concomitant therapy (ATTRACT) trial of infliximab – the first trial reported of a biological agent in RA – were found to exclude 95 % of patients seen in 2000 in the author's clinical setting [98] (Fig. 5). The three inclusion criteria were six swollen and six tender joints, met by only about one-third of patients; morning stiffness of 45 min and elevated ESR or CRP (2 of 3), met by only half the patients who had six tender and swollen joints; and a methotrexate dose greater than 12.5 mg per week which was met by only one-third of these patients. Cumulatively, these three basic inclusion criteria allowed only 5 % of RA patients to be eligible for this trial [98] (Fig. 5). Similar data have been reported in other reports [99, 100].

All clinical trials also list exclusion criteria, as many variables other than assignment to an intervention or a placebo may affect possible outcomes, such as high age, low education level, low or high disease severity, comorbidities, organ damage, fibromyalgia, previous and concomitant interventions, and many others. Exclusion criteria also restrict entry into the trial to certain possible subjects,

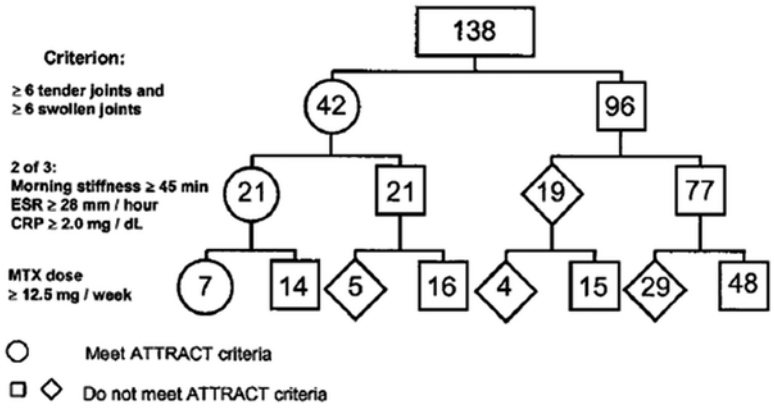

Fig. 5 Analysis of patients with rheumatoid arthritis (RA) who were potential participants in the ATTRACT (anti-tumor necrosis factor α trial in rheumatoid arthritis with concomitant therapy) trial of infliximab plus methotrexate versus methotrexate monotherapy. Of the 152 patients in this consecutive patient cohort, 12 did not have a joint count recorded and another 2 patients were taking etanercept or infliximab at the time of the first joint count and would therefore have been ineligible for the ATTRACT study. Thus, 138 patients were analyzed for meeting the inclusion criteria of the ATTRACT trial: ≥6 tender joints and ≥6 swollen joints; 2 of the following 3 – morning stiffness of ≥45 min, ESR of ≥28 mm/h, or CRP of ≥2 mg/dl; and methotrexate (MTX) dose of ≥12.5 mg/week [98] (From Sokka and Pincus [98] with permission)

in an effort to isolate observed differences to the treatment, and reduce effects of confounding variables.

In theory, the process of randomization should allow adjustment for other variables that might affect the results of a clinical trial. However, in practice, extensive exclusion criteria are common – and compromise the generalizability of results to all patients. For example, many clinicians would treat a patient with RA who is older than 80 years and has a history of breast cancer in remission for 20 years with a biological agent, the efficacy of which was documented in a clinical trial that excluded people who met both criteria for age and comorbidity.

3. *Differences between a medication and a placebo are required to be statistically significant but not necessarily robust – statistical significance may indicate only marginal clinical significance.*

A clinical trial that includes large numbers of patients may indicate that marginal clinical differences are statistically significant. For example, hundreds of clinical trials conducted during the 1970s and early 1980s indicated that various nonsteroidal anti-inflammatory drugs (NSAIDs), such as ibuprofen, naproxen, piroxicam, and diclofenac, led to improvement in RA patients in the number of tender or swollen joints.

The data were highly statistically significant in clinical trials which included relatively large numbers of patients. However, NSAIDs provided only marginal benefits to most patients [101], although a few individual patients experienced

great benefit with each of the medications. NSAIDs often are not used at all, or only on an "as needed" basis in contemporary management of RA, although efficacy was documented in dozens of clinical trials.

4. *Clinically important differences may not be statistically significant due to insufficient numbers of patients for statistical power.*
   Many important rheumatic conditions are seen in relatively small numbers by any individual health professional. Furthermore, enrollment in a clinical trial may be limited by inclusion and exclusion criteria. For example, during the early 1970s, four randomized controlled clinical trials were conducted in patients with SLE nephritis to compare mortality with combinations of prednisone plus azathioprine versus prednisone monotherapy (plus placebo). Two trials of Donadio et al. [102] and Hahn et al. [103] indicated no significant differences with combination versus prednisone monotherapy, while two others of Sztejnbok et al. [104] and Cade et al. [105] indicated lower mortality in patients treated with a combination of prednisone and azathioprine versus prednisone monotherapy (Table 4). Differences in results may be explained in part by differences in patients selected for the trials, as the two studies in which an advantage was seen to the combination included more patients with diffuse proliferative glomerulonephritis, the type of SLE nephritis with the poorest prognosis.

   Many individual trials in rheumatic diseases do not have sufficient statistical power to provide statistically significant conclusions. Such limitations in individual clinical trials in SLE nephritis have been overcome in part by a pooled analysis performed by Felson et al. (Fig. 6) [106]. The pooled analysis of eight studies indicated clear statistically significant advantages to combinations of corticosteroids plus immunosuppressive therapy versus corticosteroids alone in treatment of SLE nephritis (Fig. 6). Enhanced statistical power provided by a pooled analysis may overcome in part limitations of small numbers in individual clinical trials.

Table 4 Analysis of mortality in four randomized controlled clinical trials in SLE nephritis in which treatment with prednisone + azathioprine was compared to prednisone only

	Randomized controlled trials			
Trial characteristics/results	Sztejnbok et al. (1971) [104]	Cade et al. (1973) [105]	Donadio et al. (1974) [102]	Hahn et al. (1975) [103]
Period of observation	3 years	4 years	3 years	2 years
Prednisone monotherapy: % 4-year mortality (in *N* patients)	32 % (19)	73 % (15)	0 % (9)	30 % (13)
Prednisone + azathioprine: % 4-year mortality (in *N* patients)	0 % (16)	46 % (13)	0 % (9)	18 % (11)
Difference statistically significant?	Yes	Yes	No	No
Number with diffuse proliferative glomerulonephritis/total number	24/35	28/28	7/16	14/24

Adapted from Sztejnbok et al. [104], Cade et al. [105], Donadio et al. [102], and Hahn et al. [103]

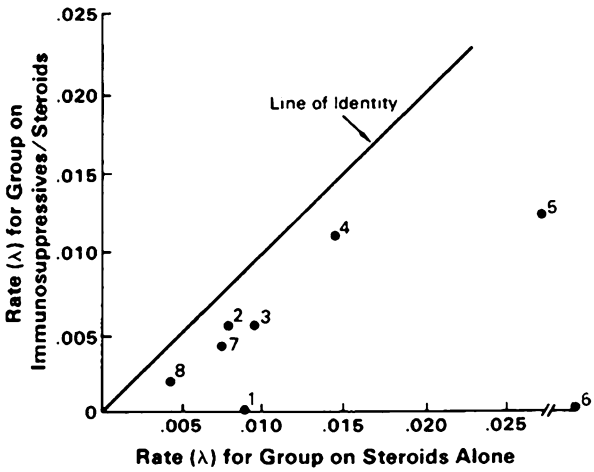

Fig. 6 Rate of renal deterioration in SLE nephritis patients treated with steroids alone and in those treated with a combination of immunosuppressive drugs and steroids. Each point represents 1 of 8 studies in a pooled analysis performed by Felson et al. [106]

Although "power calculations" are based on sophisticated mathematical computations, implying substantial precision, in actual fact, they must be based on estimates, which may involve incorrect assumptions. Furthermore, many rheumatic diseases are quite unusual, and it may be difficult to identify a sufficient number of patients to participate in the study, even if power calculations are valid. Therefore, many clinical trials may not show an effect simply because there is insufficient statistical power. Enhanced statistical power may be provided by a meta-analysis of many clinical trials, which may overcome in part limitations of small numbers in individual clinical trials [91], but even meta-analysis cannot overcome a short time frame, exclusion criteria, etc., as noted above.

5. *Important variables affecting outcomes other than whether a patient was randomized to a medication versus another medication or placebo may be seen – but usually ignored in reporting of the clinical trial.*
The basic design of the randomized controlled clinical trial is focused on identifying differences in results using one intervention versus another or a placebo, and reports of results naturally emphasize this comparison. However, in some trials, outcomes are affected more by variables other than whether a patient was randomized to a drug versus another medication or placebo.

   One example of this phenomenon is seen in the Beta-Blocker Heart Attack Trial (BHAT) study, designed to compare treatment with a beta-blocker medication, propranolol, versus placebo, to prevent death from a second heart attack in people who had suffered a recent heart attack [107]. The trial documented that propranolol was more effective than placebo (Fig. 7). However, the patients' level of formal education – a surrogate for self-management, life stress, and social support – was associated with greater differences than medication versus placebo (Fig. 7) [108]. Of note, recognition of differences according to educational level is not nearly as widely known as differences according to the medication versus placebo.

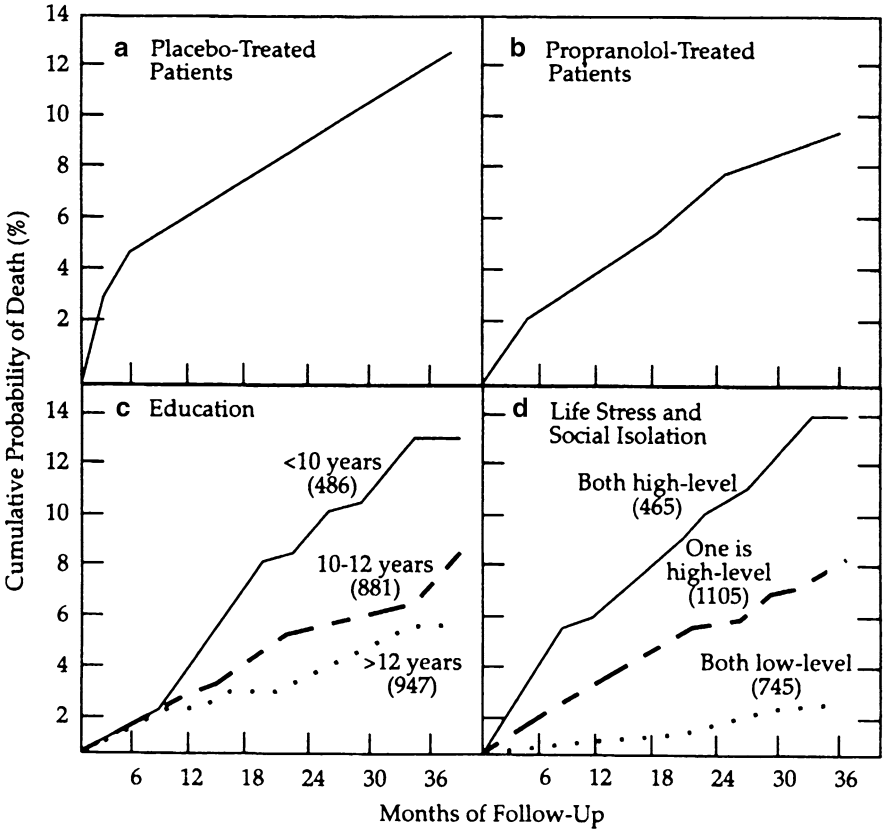

Fig. 7 Life-table cumulative mortality curves in the Beta-Blocker Heart Attack Trial (BHAT) according to (**a**) placebo treatment, (**b**) propranolol treatment, (**c**) education level, and (**d**) life stress and social isolation [108]

Another example of a clinical trial in which other variables were more significant than differences between the medication and placebo involved analysis of clofibrate versus placebo to reduce lipid levels in cardiovascular disease [109]. The 5-year mortality of patients treated with clofibrate was 20 % compared to 21 % in patients treated with placebo, a nonsignificant difference [109] (Table 5). However, 5-year mortality of patients randomized to clofibrate who adhered to their prescriptions was 15 % versus 24 % in nonadherents and virtually identical in patients randomized to placebo 15 % in adherents to placebo versus 28 % in nonadherents (P <0.0001 for adherent vs. nonadherent within each treatment arm) (Table 5). These data indicate that adherence to a treatment regimen was far more powerful to explain a reduction in mortality than whether or not patients were assigned to a lipid-lowering medication versus a placebo.

Table 5 Five-year mortality in patients given clofibrate or placebo, according to cumulative adherence to protocol prescription [109]

	Treatment group			
	Clofibrate		Placebo	
Adherence[a]	*N*	% mortality[b]	*N*	% mortality[b]
<80 %	357	24.6 ± 2.3 % (22.5 %)	882	28.2 ± 1.5 % (25.8 %)
≥80 %	708	15.0 ± 1.3 % (15.7 %)	1,813	15.1 ± 0.8 % (16.4 %)
Total study group	1,065	18.2 ± 1.2 % (18.0 %)	2,695	19.4 ± 0.8 % (19.5 %)

[a]A patient's cumulative adherence was computed as the estimated number of capsules actually taken as a percentage of the number that should have taken according to the protocol during the first 5 years of follow-up or until death (if death occurred during the first 5 years)

[b]The figures in parentheses are adjusted for 40 baseline characteristics. The figures given as percentages ±1 SE are unadjusted figures whose SEs are correcting to within 0.1 unit for the adjusted figures

6. *Traditional clinical trials with a parallel design have inflexible dosage schedules and restrict concomitant medications, although a flexible dosage schedule toward a target with multiple medications may provide optimal results.*
   Contemporary trials designed for registration of new therapies require that a new medication have statistically significantly greater efficacy than a placebo, with an acceptable profile for adverse events, in a parallel design [110]. This type of clinical trial may provide unequivocal documentation that a therapy under study is superior to placebo.

   However, this parallel design does not allow testing combinations of therapies, which are recognized increasingly as optimal for most patients with inflammatory rheumatic diseases [111]. One approach to overcome this limitation involves a "strategy trial," in which patients are treated with combinations of DMARDs versus monotherapy toward a target, generally a disease activity score – DAS [26] or DAS28 [27] – to indicate low disease activity or remission, with a protocol requiring adjustment of treatment at frequent visits.

   Eight "strategy trials" have been reported in RA [112–119] (Table 6), all of which documented significant advantages to intensification of therapies based on careful patient monitoring aimed at a target measure versus traditional therapy that was unchanged over longer periods. All eight trial results indicate that a strategy of aiming for low disease activity or remission appears more important than the agent used [120]. A "treat-to-target" strategy is emerging as the standard of care in RA [121]. Similarly, almost all patients with any inflammatory rheumatic disease are treated with combinations of medications, which cannot be studied optimally in clinical trials which restrict dosage and combinations.
7. *Surrogate markers and indices used in clinical trials may be suboptimal measures to detect changes in clinical status or predict important clinical outcomes.*
   The ultimate goal of treatment for a chronic disease is to prevent or postpone the most feared long-term consequences, such as death and disability, which generally result from poorly controlled dysregulation such as inflammation, leading to

Table 6 "Strategy" tight control clinical trials in rheumatoid arthritis

Study	Participants	Interventions	Outcomes
Pure intensive strategy versus usual care			
Grigor et al. [112]	*N* = 111, DAS >2.4, disease duration <5 years	Intensive management: monthly assessment – if DAS >2.4, escalation of therapy according to step-up protocol	*Primary*: proportion of patients with a good response (defined as a DAS <2.4 and a fall in this score from baseline by >1.2)
TICORA study		Routine care: usual rheumatology follow-up	*Secondary*: proportion of patients in remission (DAS <1.6), ACR20/50/70, radiographic progression
2 sites 18-month open-label RCT in Glasgow, Scotland		Intra-articular triamcinolone in all swollen joints	
Fransen et al. [113]	*N* = 384 meet 1987 ACR criteria	Conventional treatment	*Primary*: proportion of patients with DAS28 < 3.2 at week 24; *Secondary*: dose changes in individual DMARDs and changes in patient pain, global disease activity, and disability
Multicenter; 6-month cluster RCT at 24 sites; The Netherlands		DAS28 collected at selected visits	
Verstappen et al. [114]	*N* = 299 participants meeting the 1987 ACR criteria, disease duration <1 year	Conventional strategy	*Primary*: remission for at least 3 months – no SJC, ≤ 3TJC, ESR ≤20, global VAS ≤20
CAMERA study		Intensive strategy group according to a computer decision program	*Secondary*: improvement in single measures; mean change in disease activity
2-year multicenter open-label strategy trial			
"*Hybrid*"*: Initial parallel design treatment groups plus* "*intensive strategy*"			
Goekoop-Ruiterman et al. [115]	*N* = 508 participants meeting the 1987 ACR criteria, ≥6 SJC and TJC, disease duration ≤2 years	Sequential monotherapy	*Primary*: functional capacity by HAQ and radiographic damage by modified Sharp/van der Heijde
BeSt study		Step-up combination MTX + SSZ + HCQ	*Secondary*: ACR20/50/70 and clinical remission defined as DAS44 < 1.6
1 (2–5)-year multicenter RCT in the Netherlands		Initial combination MTX + SSZ + Prednisone	
		Initial combination MTX + infliximab	

(continued)

Table 6 (continued)

Study	Participants	Interventions	Outcomes
Hetland et al. [116]	*N* = 160 participants, disease duration <6 months	MTX + cyclosporine	*Primary*: ACR20 response at 2 years
CIMESTRA study		MTX + placebo	*Secondary*: remission, cumulative dose of betamethasone and radiographic progression
2-year multicenter placebo-controlled double-blind RCT in Denmark		Monthly assessments in both arms, betamethasone injection into all swollen joints; increase dose of MTX and/or cyclosporine by predefined protocol	
Saunders et al. [117]	*N* = 96, DAS28 > 5.1, disease duration <5 years	"Step-up" SSZ, MTX, HCQ	*Primary*: mean decrease in DAS28 at 12 months
TICORAii		Parallel triple therapy with SSZ + MTX + HCQ	*Secondary*: EULAR good responses; # in remission: ACR20/50/70
12-month RCT at 3 sites in Glasgow		Intra-articular triamcinolone in all swollen joints	
Verschueren et al. [118]; 2 years at single site in Belgium	*N* = 71 RA patients with unfavorable prognostic factors	Step-down group: modified COBRA	*Primary*: DMARD changes
		Step-up group: monotherapy with MTX, SSZ, HCQ, or AZA	*Secondary*: use of steroids, adverse events
Moreland et al. [119]	*N* = 755 meet 1987 ACR criteria, >4 TJC or SJC, disease duration <3 years	Immediate MTX-etanercept	*Primary*: change in the DAS28 between week 48 and 102
TEAR study		Immediate MTX-SSZ-HCQ	*Secondary*: radiographic progression, ACR20/50/70, modified-HAQ
2-year multicenter RCT in USA		Step-up from MTX to MTX-etanercept	
		Step-up from MTX to MTX-SSZ-HCQ	

Abbreviations: *RCT*, randomized control trial, *MTX* methotrexate *SSZ*, sulfasalazine *HCQ* hydroxychloroquine, *AZA* azathioprine, *DAS28* disease activity score for 28-joint counts, *CDAI* Clinical Disease Activity Index, *HAQ* Health Assessment Questionnaire, *TJC* tender joint count, *SJC* swollen joint count, *ESR* erythrocyte sedimentation rate

cumulative organ damage. In a clinical trial over 1 year or even 5 years unless very large numbers of patients are enrolled, it is not pragmatically possible to assess long-term outcomes such as renal or cardiac damage in hypertension or joint destruction or work disability in RA. Furthermore, damage to organs usually is irreversible by interventions designed to control the dysregulation – generally inflammation; after damage is advanced, medical interventions may be of limited to no value (only surgery, dialysis, etc., are effective).

Therefore, interventions are properly studied to analyze reversible signs of disease, which are amenable to drug therapy, such as an elevated blood pressure in hypertension, reduced CD4 counts in HIV infection, or tender or swollen

joints in RA. These "surrogate markers" are related to the long-term consequences of damage, e.g., reduced mortality rates associated with control of blood pressure [6, 7] or serum glucose [8]. However, in some instances, the correlation between surrogate markers and long-term outcomes is not robust at all. For example, there is little association between joint tenderness and radiographic damage to joints in RA [122]. Furthermore, the natural history of joint tenderness is to improve over a 5-year period, while patients may experience joint destruction with resultant deformity and limited functional capacity [123, 124]. Therefore, joint tenderness as a surrogate marker in a clinical trial may be limited in its capacity to represent future damage.

Identification of an appropriate surrogate marker has proven difficult in SLE. Indices for SLE, which are needed as no "gold standard" measure is available, have included the SLEDAI (Systemic Lupus Erythematosus Disease Activity Index) [34, 125], SLEDAI 2 K [126], BILAG (British Isles Lupus Assessment Group) index [35, 127], SLAM (Systemic Lupus Activity Measure) [128], ECLAM (European Consensus Lupus Activity Measurement) [129], and the SLICC/ACR (Systemic Lupus International Coordinating Clinics/American College of Rheumatology) damage index [130]. According to these measures, clinical trials of rituximab have shown no efficacy in SLE. By contrast, many clinicians find rituximab of value in routine clinical care of SLE [131], reminiscent of differences between clinical trial data and clinical experience with methotrexate in RA in the 1990s, noted above. One possible explanation is that the indices used to assess status and improvement in the reported clinical trials may be insufficiently sensitive to changes in SLE clinical status [132].

An example of the complexity of identifying an optimal measure for improvement in patients with SLE is seen in recent analyses of a clinical trial of abatacept in SLE [133]. The report of the clinical trial concluded that abatacept had no significant clinical efficacy in SLE. However, analyses of various clinical endpoints that have been used in other SLE clinical trials suggest that if different endpoints had been chosen, statistically significant advantages to abatacept versus placebo might have been found (Table 7) [133]. Therefore, the choice of a surrogate measure for long-term damage may greatly influence the results despite a control group.

8. *Rare adverse events cannot be identified in most trials.*
One particular limitation of clinical trials that cannot be surmounted, even with a representative sample, particularly in rare or unusual diseases such as many inflammatory rheumatic diseases, is the rare adverse event. For example, if a severe adverse event occurs in 1 in 10,000 patients, and only 1,000 are studied in clinical trials prior to approval of a medication, there is a reasonable chance that this possible important adverse event may not be observed at all in these trials. Therefore, it is probably always of value to collect post-registration surveillance data on at least 50–100,000 patients who take a given medication to monitor for unusual, but severe, adverse events. This goal can be accomplished in rheumatology if, say, 10,000 rheumatologists around the world monitored all their patients,

Table 7 Rates of complete response in patients with nephrotic levels of proteinuria (>339 mg/mmole (3 g/g)) at screening and/or baseline according to five sets of response criteria [133][a]

Criteria	Control treatment	Abatacept 10/10 treatment	Abatacept 30/10 treatment
BMS trial	1/54 (2 %)	1/49 (2 %)	2/56 (4 %)
ACR recommendations	1/54 (2 %)	3/49 (6 %)	7/56 (13 %)
LUNAR trial	2/53 (4 %)	8/48 (17 %)	13/56 (23 %)
ALMS trial	3/54 (6 %)	9/49 (18 %)	14/56 (25 %)
ACCESS trial	4/53 (8 %)	15/48 (31 %)	17/56 (30 %)

Abbreviations: *BMS* Bristol-Myers Squibb trial, *ACR* American College of Rheumatology, *LUNAR* Lupus Nephritis Assessment with Rituximab trial, *ALMS* Aspreva Lupus Management Study, *ACCESS* Abatacept and Cyclophosphamide Combination: Efficacy and Safety Study trial

[a]Patients in the abatacept treatment groups received 12 months of treatment at 10 mg/kg every 28 days (abatacept 10/10) or 12 months of treatment at 30 mg/kg every 28 days for 5 months followed by 10 mg/kg every 28 days for the remainder of the treatment period (abatacept 30/10). Values are the number of complete responders/number assessed (%)

including the 5–10 with rare diseases such as polymyositis, systemic sclerosis, and vasculitis, in identical long-term databases designed to pool the outcomes. The technology for this type of activity has been available for decades, but implementation to the rheumatology community (as well as general medical community) has been quite limited.

Intrinsic Limitations

Pragmatic limitations of clinical trials described above theoretically could be overcome by elimination of many logistical obstacles to an ideal trial. In other words, in theory, it might be possible (and desirable) to design a clinical trial that includes all patients with a given diagnosis, with no specific inclusion or exclusion criteria (other than those in whom the proposed study medication might be harmful), sufficient statistical power to observe trends that emerge from the study, indefinite continuation of the trial, and 20,000 subjects to detect rare adverse events. Nonetheless, limitations are seen to randomized controlled clinical trials that are simply intrinsic to the methodology, just as limitations exist to any scientific methodology, four of which are discussed below:

1. *The design of a clinical trial may greatly influence the results, despite inclusion of a control group.*
   The inclusion of a "control group," one of the defining characteristics in the basic design of a randomized controlled clinical trial, is commonly thought to eliminate bias in comparing results of one intervention to another or to a placebo. A control group certainly reduces bias once patients are entered into a trial, but does not eliminate all sources of bias. None the less, the design of the trial itself may strongly influence results in favor or against a particular conclusion.

Consider, for example, a simple clinical trial to compare a new medication versus placebo in a given condition. Two design options may be (*1*) to require "failure" with two previous standard treatments, so that patients have an opportunity to receive "standard of care" prior to enrollment in a clinical trial, or (*2*) to include only patients who have had no previous treatment for the condition.

Clinical research (and common sense) suggest that, in general, a medication is more likely to show efficacy when used as the first rather than as the third medication in a patient after two prior failures. A requirement for two prior failures selects for patients who are in general (although not always) more refractory to treatment. A new medication may show statistically significant differences in efficacy versus a placebo when used as the first therapy for a disease, but be only marginally better than a placebo in patients receiving the medication as their third therapy.

Another example might involve a clinical trial to compare outcomes in patients with a form of cancer who participate or do not participate in a support group. Consider two alternative designs: (*a*) offering the clinical trial to all patients at the time of diagnosis prior to any treatment; (*b*) offering the clinical trial only to patients who have persistent disease after standard treatment with surgery, radiation, and/or chemotherapy. Either design would appear quite reasonable, as patients who might be "cured" through standard treatment might be spared the trouble and expense of a support group, while "incurable" patients may benefit. However, these two designs might lead to different results. Patients beginning standard treatment would, by definition, have a higher likelihood of overall success, which might give an opportunity for a support group to add to this success. By contrast, patients who have failed standard treatment might be expected to have a lesser possibility of overall success, with a lesser possibility of additional value of a support group.

The design of a clinical trial obviously cannot preordain the results. Nonetheless, the design of a clinical trial can greatly "tilt" the probability that an intervention will or will not appear to be more efficacious than a placebo. A "control" group does not invariably eliminate biases, which are intrinsic to the design of any study.

2. *Clinical trial data are reported in groups and generally ignore individual variation*

As noted in the introductory comments, the prototype clinical trials were performed to analyze antibiotics in activity versus infectious bacteria. Bacteria are simple single-cell organisms, which present a target for antibiotics to eradicate from the body. A single optimal medication for *all* patients might be identified for treatment of a specific bacterium, particularly when the capacity of a medication to affect a target pathogen is tested in a laboratory "culture and sensitivities" analysis.

The treatment of complex multicellular and multiorgan human patients clearly is not as simple. Variation in responses among individuals to a medication would be *expected* in drugs designed to treat multicellular human organisms for such disorders as overproduction of gastric acid, control of blood pressure,

Table 8 Arthrotec compared to acetaminophen (ACTA) crossover clinical trial: patient ratings of each drug [134]

Patient ratings	Group I Arthrotec → acetaminophen *N* (%)	Group II acetaminophen → Arthrotec *N* (%)	Total *N* (%)
Arthrotec better or much better	52 (58 %)	48 (57 %)	100 (57 %)
No difference	18 (20 %)	21 (25 %)	39 (22 %)
Acetaminophen better or much better	20 (22 %)	15 (18 %)	35 (20 %)

management of pain, or reduction of depression. However, in general, a clinical trial is reported to identify the "best" therapy for *all* patients, rather than to identify *which* therapy might be most effective for *particular* individual patients.

Crossover clinical trials in which two agents are compared often document the phenomenon of individual variation in responses to two or more treatments. For example, in a crossover clinical trial of Arthrotec (diclofenac coated with misoprostol) compared to acetaminophen (ACTA trial), 57 % of individuals reported that Arthrotec was superior, 21 % that acetaminophen was superior, and 22 % found the two medications of equal efficacy (Table 8) [134]. An accurate interpretation of the data might be that Arthrotec is better for most individuals, but acetaminophen is better for some individual patients. However, the usual interpretation of such data is that Arthrotec is "superior" to acetaminophen in general.

This limitation may have significant consequences for therapies for individual patients. Many hospital formularies will select only a single medication from a given category, such as H2 blockers to reduce acid in peptic ulcer and reflux disease, tricyclic antidepressants, or NSAIDs. The reasoning is that it is cost-effective to have available only a single medication among several that all act according to a similar mechanism. However, each individual medication may be superior in some individual patients, due to different receptors, metabolism, and other idiosyncratic characteristics of the host. The interpretation that a single optimal medication exists for *all* patients with a disease is an incorrect assumption, probably based in large part on the origin of clinical trials in studies of antibiotic medications designed to interact primarily with simple bacterial cells, rather than complex human organisms.

3. *Interpretation of adverse events is not standardized and depends on assessment of risks and benefits which differ widely among individuals.*
   All interventions, including medications, physical therapy, exercise programs, etc., may be associated with some type of adverse event in certain individuals, ranging from renal damage to inconvenient travel to a support group. Consider, for example, in a comparison of two medications, that Medication A leads to remission in 95 % of patients with few adverse event effects, but 1 in 10,000

patients experiences renal failure, while Medication B leads to improvement in 50 % of patients, no remissions, and "nuisance" gastrointestinal side effects in 20 % of patients, but leads to no severe harm to internal organs. Which is the preferred medication? That depends in large part on how an individual patient assesses risks and benefits, which varies widely among individual patients and individuals in general. As an example to patients, the author often asks a patient to assess the risk/benefit of playing a lottery, pointing out that there is no single "correct" answer. A committee verdict concerning groups cannot provide optimal guidance to each individual patient.

In general, the community of health professionals accepts the interpretation of the authors of a clinical trial regarding risks and benefits of a therapy. However, in actual practice, certain patients may prefer the odds of one alternative or the other and can make an informed decision as to which is the optimal treatment. A clinical trial infers a "black and white" choice, while the actual results suggest "shades of gray." This interpretive component in analysis of results of a clinical trial may explain occasional contentious disagreement within FDA advisory groups concerning approval of certain new medications or procedures such as mammography in 40- to 50-year-old women. A positive or negative recommendation depends on analysis of risks versus benefits, interpretation of which varies greatly among patients (as well as "experts").

4. *The format of a clinical trial compromises the "placebo effect" in not informing patients that they may not receive the "best" therapy.*
Considerable information has been reported over the last few decades concerning the "placebo" effect in any patient intervention [135]. After all, until the twentieth century, most medications were of little efficacy and yet health professionals were highly regarded as providing "curative" medications in many situations. This placebo effect is compromised considerably when a health professional invites a patient to participate in a "scientific experiment" to recognize the best therapy, rather than telling a patient that she/he will receive the "best therapy."

Most clinical trials show substantial benefit to participants in both placebo and treatment groups, suggesting that there nonetheless exists a considerable "placebo effect" even within the clinical trial methodology. It may be argued that both arms of a clinical trial are diminished in their therapeutic efficacy by loss of possible placebo effect, but this loss is "controlled for." It seems clear, nonetheless, that some of the therapeutic "placebo" benefit which results from patients being told that they are being given an optimal therapy is lost in the circumstances of the clinical trial.

Summary and Conclusion

Clinical trials remain the optimal method to compare one therapy with another or a placebo independent of inevitable biases associated with choices of therapies [1, 82]. Nonetheless, clinical trials have limitations. In this chapter, selected randomized controlled clinical trials conducted in chronic rheumatic diseases, including

RA, SLE, polymyositis, and OA, as well as other chronic cardiovascular diseases, have been summarized to illustrate some of these limitations. Some limitations may be overcome by longer trials and meta-analyses. However, pragmatic and intrinsic limitations will always affect the clinical trial methodology to some extent. A greater awareness of these limitations would be of benefit to health professionals and the general public in interpreting results and implications of clinical trials for clinical care.

References

1. Feinstein AR. An additional basic science for clinical medicine: II. The limitations of randomized trials. Ann Intern Med. 1983;99:544–50.
2. Engel GL. The biopsychosocial model and the education of health professionals. Ann N Y Acad Sci. 1978;310:169–81.
3. Howick J. The philosophy of evidence-based medicine. Oxford: Wiley-Blackwell; 2011.
4. Streptomycin treatment of pulmonary tuberculosis. Br Med J. 1948;2(4582):769–82.
5. Hill AB. Suspended judgment: memories of the British streptomycin trial in tuberculosis. The first randomized clinical trial. Control Clin Trials. 1990;11:77–9.
6. Veterans Administration Cooperative Study on Antihypertensive Agent. Effects of treatment on morbidity in hypertension: results in patients with diastolic blood pressures averaging 115 through 129 mm Hg. JAMA. 1967;202:1028–34.
7. 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.
8. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group. N Engl J Med. 1993;329(14):977–86.
9. Krause D, Schleusser B, Herborn G, Rau R. Response to methotrexate treatment is associated with reduced mortality in patients with severe rheumatoid arthritis. Arthritis Rheum. 2000;43:14–21.
10. Choi HK, Hernán MA, Seeger JD, Robins JM, Wolfe F. Methotrexate and mortality in patients with rheumatoid arthritis: a prospective study. Lancet. 2002;359:1173–7.
11. Wolfe F, Michaud K, Gefeller O, Choi HK. Predicting mortality in patients with rheumatoid arthritis. Arthritis Rheum. 2003;48:1530–42.
12. Jacobsson LTH, Turesson C, Nilsson JA, Petersson IF, Lindqvist E, Saxne T, et al. Treatment with TNF blockers and mortality risk in patients with rheumatoid arthritis. Ann Rheum Dis. 2007;66(5):670–5.
13. Pincus T, Gibofsky A, Weinblatt ME. Urgent care and tight control of rheumatoid arthritis as in diabetes and hypertension: better treatments but a shortage of rheumatologists. Arthritis Rheum. 2002;46(4):851–4.
14. Pincus T, Yazici Y, Sokka T. Complexities in assessment of rheumatoid arthritis: absence of a single gold standard measure. Rheum Dis Clin North Am. 2009;35(4):687–97.
15. Waaler E. On the occurrence of a factor in human serum activating the specific agglutination of sheep blood corpuscles. APMIS. 1940;17:172–8.
16. Rose HM, Ragan C, Pearce E, Lipman MO. Differential agglutination of normal and sensitized sheep erythrocytes by sera of patients with rheumatoid arthritis. Proc Soc Exp Biol Med. 1948;68:1–6.
17. Hargraves MM, Richmond H, Morton R. Presentation of two bone marrow elements: the "tart" cell and "L.E." cell. Proc Staff Meet Mayo Clin. 1948;23:25–8.

18. Nishimura K, Sugiyama D, Kogata Y, Tsuji G, Nakazawa T, Kawano S, et al. Meta-analysis: diagnostic accuracy of anti-cyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis. Ann Intern Med. 2007;146(11):797–808.
19. Wolfe F, Michaud K. The clinical and research significance of the erythrocyte sedimentation rate. J Rheumatol. 1994;21:1227–37.
20. Wolfe F. Comparative usefulness of c-reactive protein and erythrocyte sedimentation rate in patients with rheumatoid arthritis. J Rheumatol. 1997;24:1477–85.
21. Sokka T, Pincus T. Erythrocyte sedimentation rate, C-reactive protein, or rheumatoid factor are normal at presentation in 35 %-45% of patients with rheumatoid arthritis seen between 1980 and 2004: analyses from Finland and the United States. J Rheumatol. 2009;36(7): 1387–90.
22. Munves EF, Schur PH. Antibodies to Sm and RNP: prognosticators of disease involvement. Arthritis Rheum. 1983;26:848–53.
23. Pincus T. A pragmatic approach to cost-effective use of laboratory tests and imaging procedures in patients with musculoskeletal symptoms. Prim Care. 1993;20:795–814.
24. Pincus T. Laboratory tests in rheumatic disorders. In: Klippel JH, Dieppe PA, editors. Rheumatology. 2nd ed. London: Mosby International; 1997. p. 10.1–8.
25. Goldsmith CH, Smythe HA, Helewa A. Interpretation and power of pooled index. J Rheumatol. 1993;20:575–8.
26. van der Heijde DM, van't Hof M, van Riel PL, van de Putte LB. Development of a disease activity score based on judgment in clinical practice by rheumatologists. J Rheumatol. 1993;20:579–81.
27. Prevoo MLL, van't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LBA, van Riel PLCM. Modified disease activity scores that include twenty-eight-joint counts: development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum. 1995;38:44–8.
28. Aletaha D, Smolen J. The simplified disease activity index (SDAI) and the clinical disease activity index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. Clin Exp Rheumatol. 2005;23:S100–8.
29. Pincus T, Strand V, Koch G, Amara I, Crawford B, Wolfe F, et al. An index of the three core data set patient questionnaire measures distinguishes efficacy of active treatment from placebo as effectively as the American College of Rheumatology 20 % response criteria (ACR20) or the disease activity score (DAS) in a rheumatoid arthritis clinical trial. Arthritis Rheum. 2003;48(3):625–30.
30. Pincus T, Bergman MJ, Yazici Y, Hines P, Raghupathi K, Maclean R. An index of only patient-reported outcome measures, routine assessment of patient index data 3 (RAPID3), in two abatacept clinical trials: similar results to disease activity score (DAS28) and other RAPID indices that include physician-reported measures. Rheumatology (Oxford). 2008; 47(3):345–9.
31. Pincus T, Swearingen CJ, Bergman M, Yazici Y. RAPID3 (routine assessment of patient index data 3), a rheumatoid arthritis index without formal joint counts for routine care: proposed severity categories compared to DAS and CDAI categories. J Rheumatol. 2008;35: 2136–47.
32. Petri M, Hellmann DB, Hochberg M. Validity and reliability of lupus activity measures in the routine clinic setting. J Rheumatol. 1992;19:53–9.
33. Bencivelli W, Vitali C, Isenberg DA, Smolen JS, Snaith ML, Sciuto M, et al. Disease activity in systemic lupus erythematosus: report of the Consensus Study Group of the European Workshop for Rheumatology Research III. Development of a computerised clinical chart and its application to the comparison of different indices of disease activity. The European Consensus Study Group for Disease Activity in SLE. Clin Exp Rheumatol. 1992;10(5): 549–54.
34. Hawker G, Gabriel S, Bombardier C, Goldsmith C, Caron D, Gladman D. A reliability study of SLEDAI: a disease activity index for systemic lupus erythematosus. J Rheumatol. 1993;20: 657–60.

35. Hay EM, Bacon PA, Gordon C, Isenberg DA, Maddison P, Snaith ML, et al. The BILAG index: a reliable and valid instrument for measuring clinical disease activity in systemic lupus erythematosus. Q J Med. 1993;86(7):447–58.
36. Gladman DD, Urowitz MB, Goldsmith CH, Fortin P, Ginzler E, Gordon C, et al. The reliability of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index in patients with systemic lupus erythematosus. Arthritis Rheum. 1997;40(5):809–13.
37. Mosca M, Bencivelli W, Vitali C, Carrai P, Neri R, Bombardieri S. The validity of the ECLAM index for the retrospective evaluation of disease activity in systemic lupus erythematosus. Lupus. 2000;9(6):445–50.
38. Swaak AJ, van den Brink HG, Smeenk RJ, Manger K, Kalden JR, Tosi S, et al. Systemic lupus erythematosus. Disease outcome in patients with a disease duration of at least 10 years: second evaluation. Lupus. 2001;10(1):51–8.
39. Lam GKW, Petri M. Assessment of systemic lupus erythematosus. Clin Exp Rheumatol. 2005;23:S120–32.
40. Luqmani RA, Bacon PA, Moots RJ, Janssen BA, Pall A, Emery P, et al. Birmingham Vasculitis Activity Score (BVAS) in systemic necrotizing vasculitis. Q J Med. 1994;87:671–8.
41. Bacon PA, Moots RJ, Exley E, Luqmani R, Rasmussen N. VITAL assessment of vasculitis – workshop report. Clin Exp Rheumatol. 1995;13:275–8.
42. Exley AR, Bacon PA, Luqmani RA, Kitas GD, Gordon C, Savage COS, et al. Development and initial validation of the Vasculitis Damage Index for the standardized clinical assessment of damage in the systemic vasculitides. Arthritis Rheum. 1997;40:371–80.
43. Whiting O'Keefe QE, Stone JH, Hellmann DB. Validity of a vasculitis activity index for systemic necrotizing vasculitis. Arthritis Rheum. 1999;42:2365–71.
44. Stone JH, Hoffman GS, Merkel PA, Min Y, Uhlfelder ML, Hellmann DB, et al. A disease-specific activity index for Wegener's granulomatosis: modification of the Birmingham Vasculitis Activity Score. Arthritis Rheum. 2001;44:912–20.
45. Seo P, Min Y, Holbrook JT, Hoffman GS, Merkel PA, Spiera R, et al. Damage caused by Wegener's granulomatosis and its treatment: prospective data from the Wegener's Granulomatosis Etanercept Trail (WGET). Arthritis Rheum. 2005;52:2168–78.
46. Clegg DO, Reda DJ, Weisman MH, Blackburn WD, Cush JJ, Cannon GW, et al. Comparison of sulfasalazine and placebo in the treatment of ankylosing spondylitis. A Department of Veterans Affairs Cooperative Study. Arthritis Rheum. 1996;39(12):2004–12.
47. Fleischer JAB, Feldman SR, Rapp SR, Reboussin DM, Exum ML, Clark AR, et al. Disease severity measures in a population of psoriasis patients: the symptoms of psoriasis correlate with self-administered psoriasis area severity index scores. J Invest Dermatol. 1996;107: 26–9.
48. Kavanaugh A, Cassell S. The assessment of disease activity and outcomes in psoriatic arthritis. Clin Exp Rheumatol. 2005;23:S142–7.
49. Dougados M, Gueguen A, Nakache JP, Nguyen M, Mery C, Amor B. Evaluation of a functional index and an articular index in ankylosing spondylitis. J Rheumatol. 1988;15:302–7.
50. Calin A, Garrett S, Whitelock H, Kennedy LG, O'Hea J, Mallorie P, et al. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. J Rheumatol. 1994;21(12):2281–5.
51. Calin A, Nakache JP, Gueguen A, Zeidler H, Mielants H, Dougados M. Defining disease activity in ankylosing spondylitis: is a combination of variables (Bath Ankylosing Spondylitis Disease Activity Index) an appropriate instrument? Rheumatology (Oxford). 1999;38(9): 878–82.
52. Calin A, MacKay K, Santos H, Brophy S. A new dimension to outcome: application of the Bath Ankylosing Spondylitis Radiology Index. J Rheumatol. 1999;26(4):988–92.
53. Zochling J, Braun J. Assessment of ankylosing spondylitis. Clin Exp Rheumatol. 2005;23: S133–41.
54. Pincus T, Segurado OG. Most visits of most patients with rheumatoid arthritis to most rheumatologists do not include a formal quantitative joint count. Ann Rheum Dis. 2006;65: 820–2.

55. Anderson J, Caplan L, Yazdany J, Robbins ML, Neogi T, Michaud K, et al. Rheumatoid arthritis disease activity measures: American College of Rheumatology recommendations for use in clinical practice. Arthritis Care Res (Hoboken). 2012;64(5):640–7.
56. Castrejón I, McCollum L, Durusu Tanriover M, Pincus T. Importance of patient history and physical examination in rheumatoid arthritis compared to other chronic diseases: results of a physician survey. Arthritis Care Res. 2012;64(8):1250–5.
57. Pincus T, Swearingen CJ. The HAQ compared with the MDHAQ: "keep it simple, stupid" (KISS), with feasibility and clinical value as primary criteria for patient questionnaires in usual clinical care. Rheum Dis Clin North Am. 2009;35(4):787–98.
58. Pincus T, Skummer PT, Grisanti MT, Castrejón I, Yazici Y. MDHAQ/RAPID3 can provide a roadmap or agenda for all rheumatology visits when the entire MDHAQ is completed at all patient visits and reviewed by the doctor before the encounter. Bull NYU Hosp Jt Dis. 2012;70(3):177–86.
59. Pincus T, Oliver AM, Bergman MJ. How to collect an MDHAQ to provide rheumatology vital signs (function, pain, global status, and RAPID3 scores) in the infrastructure of rheumatology care, including some misconceptions regarding the MDHAQ. Rheum Dis Clin North Am. 2009;35(4):799–812.
60. OCEBM Levels of Evidence Working Group. The Oxford 2011 Levels of Evidence. http://www.cebm.net/index.aspx?o=5653, Oxford Centre for Evidence-Based Medicine. Oxford: Oxford Centre for Evidence-Based Medicine; 2011. Ref Type: Generic.
61. Smith GCS, Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. BMJ. 2003;327:1459–61.
62. Glasziou P, Chalmers I, Rawlins M, McCulloch P. When are randomised trials unnecessary? Picking signal from noise. BMJ. 2007;334(7589):349–51.
63. Aronson JK, Hauben M. Anecdotes that provide definitive evidence. BMJ. 2006;333(7581):1267–9.
64. Freiman JA, Chalmers TC, Smith Jr H, Kuebler RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial: survey of 71 "negative" trials. N Engl J Med. 1978;299:690–4.
65. Sackett DL, Gent M. Controversy in counting and attributing events in clinical trials. N Engl J Med. 1979;301:1410–2.
66. Sackett DL. The competing objectives of randomized trials. N Engl J Med. 1980;303:1059–60.
67. Huskisson EC. Important factors in the success and failure of clinical trials (closing remarks). Agents Actions Suppl. 1980;7:323–4.
68. Freireich EJ. The randomized clinical trial as an obstacle to clinical research. In: Varco RL, Delaney JP, editors. Controversy in surgery. 2nd ed. Philadelphia: W.B. Saunders; 1983. p. 5–12.
69. Diamond GA, Forrester JS. Clinical trials and statistical verdicts: probable grounds for appeal. Ann Intern Med. 1983;98:385–94.
70. Chalmers TC, Celano P, Sacks HS, Smith Jr H. Bias in treatment assignment in controlled clinical trials. N Engl J Med. 1983;309:1358–61.
71. Bombardier C, Tugwell P. Controversies in the analysis of longterm clinical trials of slow acting drugs (editorial). J Rheumatol. 1985;12:403–5.
72. Guyatt G, Sackett D, Taylor DW, Chong J, Roberts R, Pugsley S. Determining optimal therapy – randomized trials in individual patients. N Engl J Med. 1986;314:889–92.
73. Gotzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. Control Clin Trials. 1989;10:31–56.
74. Sanz I, Dang H, Takei M, Talal N, Capra JD. VH sequence of a human anti-Sm autoantibody. Evidence that autoantibodies can be un-mutated copies of germline genes. J Immunol. 1989;142:883–7.
75. Klippel JH. Comment: winning the battle, losing the war? Another editorial about RA. J Rheumatol. 1990;17:1118–22.

76. Felson DT, Anderson JJ, Meenan RF. Time for changes in the design, analysis, and reporting of rheumatoid arthritis clinical trials. Arthritis Rheum. 1990;33:140–9.
77. Hawley DJ, Wolfe F. Are the results of controlled clinical trials and observational studies of second line therapy in rheumatoid arthritis valid and generalizable as measures of rheumatoid arthritis outcome: analysis of 122 studies. J Rheumatol. 1991;18:1008–14.
78. Rothwell PM. Can overall results of clinical trials be applied to all patients? Lancet. 1995;345:1616–9.
79. Cleophas TJ, Zwinderman AH. Limitations of randomized clinical trials. Proposed alternative designs. Clin Chem Lab Med. 2000;38(12):1217–23.
80. Grossman J, Mackenzie FJ. The randomized controlled trial: gold standard, or merely standard? Perspect Biol Med. 2005;48(4):516–34.
81. Ho PM, Peterson PN, Masoudi FA. Evaluating the evidence: is there a rigid hierarchy? Circulation. 2008;118(16):1675–84.
82. Pincus T. Rheumatoid arthritis: disappointing long-term outcomes despite successful short-term clinical trials. J Clin Epidemiol. 1988;41:1037–41.
83. Pincus T, Wolfe F. Response to letter: gold therapy for rheumatoid arthritis: challenges to traditional paradigms. Ann Intern Med. 1992;117:169–70.
84. Pincus T. Limitations of randomized controlled clinical trials to recognize possible advantages of combination therapies in rheumatic diseases. Semin Arthritis Rheum. 1993;23 Suppl 1:2–10.
85. Pincus T, Stein M. What is the best source of useful data on the treatment of rheumatoid arthritis: clinical trials, clinical observations, or clinical protocols? J Rheumatol. 1995;22: 1611–7.
86. Pincus T, Stein CM. Why randomized controlled clinical trials do not depict accurately long-term outcomes in rheumatoid arthritis: some explanations and suggestions for future studies. Clin Exp Rheumatol. 1997;15 Suppl 17:S27–38.
87. Pincus T, Sokka T. Clinical trials in rheumatic diseases: designs and limitations. Rheum Dis Clin North Am. 2004;30:701–4.
88. Pincus T, Sokka T. Complexities in the quantitative assessment of patients with rheumatic diseases in clinical trials and clinical care. Clin Exp Rheumatol. 2005;23:S1–9.
89. Chalmers TC, Block JB, Lee S. Controlled studies in clinical cancer research. N Engl J Med. 1972;287:75–8.
90. Williams HJ, Ward JR, Reading JC, Brooks RH, Clegg DO, Skosey JL, et al. Comparison of auranofin, methotrexate, and the combination of both in the treatment of rheumatoid arthritis: a controlled clinical trial. Arthritis Rheum. 1992;35:259–69.
91. Felson DT, Anderson JJ, Meenan RF. The comparative efficacy and toxicity of second-line drugs in rheumatoid arthritis: results of two metaanalyses. Arthritis Rheum. 1990;33:1449–61.
92. Pincus T, Marcum SB, Callahan LF. Long-term drug therapy for rheumatoid arthritis in seven rheumatology private practices: II. Second-line drugs and prednisone. J Rheumatol. 1992;19: 1885–94.
93. Donahue KE, Gartlehner G, Jonas DE, Lux LJ, Thieda P, Jonas BL, et al. Systematic review: comparative effectiveness and harms of disease-modifying medications for rheumatoid arthritis. Ann Intern Med. 2008;148(2):124–34.
94. Sokka T, Kautiainen H, Toloza S, Makinen H, Verstappen SM, Hetland ML, et al. QUEST-RA: quantitative clinical assessment of patients with rheumatoid arthritis seen in standard rheumatology care in 15 countries. Ann Rheum Dis. 2007;66:1491–6.
95. Bunch TW, Worthington JW, Combs JJ, Ilstrup DM, Engel AG. Azathioprine with prednisone for polymyositis: a controlled, clinical trial. Ann Intern Med. 1980;92:365–9.
96. Bunch TW. Prednisone and azathioprine for polymyositis: long-term followup. Arthritis Rheum. 1981;24:45–8.
97. Austin III HA, Klippel JH, Balow JE, le Riche NGH, Steinberg AD, Plotz PH, et al. Therapy of lupus nephritis: controlled trial of prednisone and cytotoxic drugs. N Engl J Med. 1986; 314:614–9.
98. Sokka T, Pincus T. Eligibility of patients in routine care for major clinical trials of anti-tumor necrosis factor alpha agents in rheumatoid arthritis. Arthritis Rheum. 2003;48(2):313–8.

99. Sokka T, Pincus T. An early rheumatoid arthritis treatment evaluation registry (ERATER) in the United States. Clin Exp Rheumatol. 2005;23:S178–81.
100. Gogus F, Yazici Y, Yazici H. Inclusion criteria as widely used for rheumatoid arthritis clinical trials: patient eligibility in a Turkish cohort. Clin Exp Rheumatol. 2005;23(5):681–4.
101. Pincus T, Marcum SB, Callahan LF, Adams RF, Barber J, Barth WF, et al. Long-term drug therapy for rheumatoid arthritis in seven rheumatology private practices: I. Nonsteroidal anti-inflammatory drugs. J Rheumatol. 1992;19:1874–84.
102. Donadio Jr JV, Holley KE, Wagoner RD, Ferguson RH, McDuffie FC. Further observations on the treatment of lupus nephritis with prednisone and combined prednisone and azathioprine. Arthritis Rheum. 1974;17:573–81.
103. Hahn BH, Kantor OS, Osterland CK. Azathioprine plus prednisone compared with prednisone alone in the treatment of systemic lupus erythematosus: report of a prospective controlled trial in 24 patients. Ann Intern Med. 1975;83:597–605.
104. Sztejnbok M, Stewart A, Diamond H, Kaplan D. Azathioprine in the treatment of systemic lupus erythematosus. Arthritis Rheum. 1971;14:639–45.
105. Cade R, Spooner G, Schlein E, Pickering M, DeQuesada A, Holcomb A, et al. Comparison of azathioprine, prednisone, and heparin alone or combined in treating lupus nephritis. Nephron. 1973;10:37–56.
106. Felson DT, Anderson J. Evidence for the superiority of immunosuppressive drugs and prednisone over prednisone alone in lupus nephritis: results of a pooled analysis. N Engl J Med. 1984;311:1528–33.
107. A randomized trial of propranolol in patients with acute myocardial infarction. I. Mortality results. JAMA. 1982;247(12):1707–14.
108. Ruberman W, Weinblatt E, Goldberg JD, Chaudhary BS. Psychosocial influences on mortality after myocardial infarction. N Engl J Med. 1984;311:552–9.
109. The Coronary Drug Project Research Group. Influence of adherence to treatment and response of cholesterol on mortality in the coronary drug project. N Engl J Med. 1980;303:1038–41.
110. Boers M. Add-on or step-up trials for new drug development in rheumatoid arthritis: a new standard? Arthritis Rheum. 2003;48:1481–3.
111. Smolen JS, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas D, Burmester G, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis. 2010;69(4):631–7.
112. Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. Lancet. 2004;364:263–9.
113. Fransen J, Moens HB, Speyer I, van Riel PL. Effectiveness of systematic monitoring of rheumatoid arthritis disease activity in daily practice: a multicentre, cluster randomised controlled trial. Ann Rheum Dis. 2005;64:1294–8.
114. Verstappen SM, Jacobs JW, van der Veen MJ, Heurkens AH, Schenk Y, Ter Borg EJ, et al. Intensive treatment with methotrexate in early rheumatoid arthritis: aiming for remission. Computer Assisted Management in Early Rheumatoid Arthritis (CAMERA, an open-label strategy trial). Ann Rheum Dis. 2007;66(11):1443–9.
115. Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, van Zeben D, Kerstens PJ, Hazes JM, et al. Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): a randomized, controlled trial. Arthritis Rheum. 2005;52:3381–90.
116. Hetland ML, Stengaard-Pedersen K, Junker P, Lottenburger T, Hansen I, Andersen LS, et al. Aggressive combination therapy with intra-articular glucocorticoid injections and conventional disease-modifying anti-rheumatic drugs in early rheumatoid arthritis: second-year clinical and radiographic results from the CIMESTRA study. Ann Rheum Dis. 2008;67:815–22.
117. Saunders SA, Capell HA, Stirling A, Vallance R, Kincaid W, McMahon AD, et al. Triple therapy in early active rheumatoid arthritis: a randomized, single-blind, controlled trial comparing step-up and parallel treatment strategies. Arthritis Rheum. 2008;58(5):1310–7.

118. Verschueren P, Esselens G, Westhovens R. Daily practice effectiveness of a step-down treatment in comparison with a tight step-up for early rheumatoid arthritis. Rheumatology (Oxford). 2008;47(1):59–64.
119. Moreland LW, O'Dell JR, Paulus HE, Curtis JR, Bathon JM, St Clair EW, et al. A randomized comparative effectiveness study of oral triple therapy versus etanercept plus methotrexate in early aggressive rheumatoid arthritis: the treatment of Early Aggressive Rheumatoid Arthritis Trial. Arthritis Rheum. 2012;64(9):2824–35.
120. Sokka T, Pincus T. Rheumatoid arthritis: strategy more important than agent. Lancet. 2009;374(9688):430–2.
121. Turchetti G, Smolen JS, Kavanaugh A, Braun J, Pincus T. Treat-to-target in rheumatoid arthritis: clinical and pharmacoeconomic considerations. Clin Exp Rheumatol. 2012;30(4 Suppl 73):S1–169. Ref Type: Journal (Full).
122. Fuchs HA, Callahan LF, Kaye JJ, Brooks RH, Nance EP, Pincus T. Radiographic and joint count findings of the hand in rheumatoid arthritis: related and unrelated findings. Arthritis Rheum. 1988;31:44–51.
123. Callahan LF, Pincus T, Huston III JW, Brooks RH, Nance Jr EP, Kaye JJ. Measures of activity and damage in rheumatoid arthritis: depiction of changes and prediction of mortality over five years. Arthritis Care Res. 1997;10:381–94.
124. Mulherin D, Fitzgerald O, Bresnihan B. Clinical improvement and radiological deterioration in rheumatoid arthritis: evidence that pathogenesis of synovial inflammation and articular erosion may differ. Br J Rheumatol. 1996;35:1263–8.
125. Bombardier C, Gladman DD, Urowitz MB, Caron D, Chang CH. Derivation of the SLEDAI A disease activity index for lupus patients. The Committee on Prognosis Studies in SLE. Arthritis Rheum. 1992;35(6):630–40.
126. Gladman DD, Ibanez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. J Rheumatol. 2002;29(2):288–91.
127. Isenberg DA, Rahman A, Allen E, Farewell V, Akil M, Bruce IN, et al. BILAG 2004. Development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. Rheumatology (Oxford). 2005;44(7):902–6.
128. Liang MH, Socher SA, Larson MG, Schur PH. Reliability and validity of six systems for the clinical assessment of disease activity in systemic lupus erythematosus. Arthritis Rheum. 1989;32:1107–18.
129. Vitali C, Bencivelli W, Isenberg DA, Smolen JS, Snaith ML, Sciuto M, et al. Disease activity in systemic lupus erythematosus: report of the Consensus Study Group of the European Workshop for Rheumatology Research II. Identification of the variables indicative of disease activity and their use in the development of an activity score. The European Consensus Study Group for Disease Activity in SLE. Clin Exp Rheumatol. 1992;10(5):541–7.
130. Gladman D, Ginzler E, Goldsmith C, Fortin P, Liang M, Urowitz M, et al. The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index for systemic lupus erythematosus. Arthritis Rheum. 1996;39:363–9.
131. Gunnarsson I, Jonsdottir T. Rituximab treatment in lupus nephritis–where do we stand? Lupus. 2013;22(4):381–9.
132. Petri M. Disease activity assessment in SLE: do we have the right instruments? Ann Rheum Dis. 2007;66 Suppl 3:iii61–4.
133. Wofsy D, Hillson JL, Diamond B. Abatacept for lupus nephritis: alternative definitions of complete response support conflicting conclusions. Arthritis Rheum. 2012;64(11):3660–5.
134. Pincus T, Koch GG, Sokka T, Lefkowith J, Wolfe F, Jordan JM, et al. A randomized, double-blind, crossover clinical trial of diclofenac plus misoprostol versus acetaminophen in patients with osteoarthritis of the hip or knee. Arthritis Rheum. 2001;44:1587–98.
135. Kaptchuk TJ. Powerful placebo: the dark side of the randomised controlled trial. Lancet. 1998;351:1722–5.