Limitations of clinical trials in chronic diseases: is the efficacy of methotrexate (MTX) underestimated in polyarticular psoriatic arthritis on the basis of limitations of clinical trials more than on limitations of MTX, as was seen in rheumatoid arthritis?

T. Pincus¹, M.J. Bergman², Y. Yazici³

¹Department of Internal Medicine, Division of Rheumatology, Rush University Medical Center, Chicago, USA; ²Arthritis and Rheumatology, Taylor Hospital, Ridley Park, PA, USA; ³Division of Rheumatology, Department of Internal Medicine, NYU Hospital for Joint Diseases, New York, USA.

Theodore Pincus, MD
Martin J. Bergman, MD
Yusuf Yazici, MD

Please address correspondence to: Theodore Pincus, MD, Division of Rheumatology, Rush University Medical Center, 1611 West Harrison Street, Suite 510, Chicago, IL 60612, USA; E-mail: tedpincus@gmail.com

Received and accepted on September 30, 2015.
Clin Exp Rheumatol 2015; 33 (Suppl. 93): S82-S93.

© Copyright CLINICAL AND EXPERIMENTAL RHEUMATOLOGY 2015.

Key words: clinical trials, psoriatic arthritis, rheumatoid arthritis, methotrexate

T. Pincus is president of Health Report Services, Inc., which receives royalties and license fees from for-profit organisations, including pharmaceutical companies and electronic medical record companies, for use of MDHAQ/RAPID3, based on a copyright and trademark. No license is collected from clinicians and non-profit researchers, who may freely use paper copies of MDHAQ/RAPID3 to monitor patient status in usual clinical care and academic research. All license fees are used to support further development of quantitative measurement in clinical rheumatology care;

M.J. Bergman is a consultant/speaker for Amgen, Genentech, Celgene, Abbvie, Janssen, Pfizer and a shareholder for Merck, Pfizer and Janssen (through Johnson and Johnson);

Y. Yazici receives consulting fees from Bristol-Myers-Squibb and Celgene, and research support from Bristol-Myers-Squibb, Celgene, and Genentech.

ABSTRACT
Clinical trials are the optimal method to establish efficacy of a drug versus placebo or another drug. Nonetheless, important limitations are seen, particularly in chronic diseases over long periods, although most are ignored. Pragmatic limitations of clinical trials include a relatively short observation period, suboptimal dosage schedules, suboptimal surrogate markers for long-term outcomes, statistically significant results which may not be clinically unimportant and vice versa. Even ideal clinical trials have intrinsic limitations, including the influence of design on results, data reported in groups which ignore individual variation, non-standard observer-dependent interpretation of a balance of efficacy and toxicity, and distortion of a “placebo effect.” Limitations are seen in many clinical trials of methotrexate (MTX) in rheumatoid arthritis (RA) and psoriatic arthritis (PsA). The first MTX clinical trial in rheumatology documented excellent efficacy in PsA, but frequent adverse events in 1964, explained by intravenous doses up to 150 kg. MTX was abandoned until the 1980s for RA, while gold salts and penicillamine were termed “remission-inducing,” on the basis limitations of clinical trials. In the most recent MTX in PsA (MIP A) trial, all outcomes favoured MTX, but only patient and physician global estimates met the p<0.05 criterion. A conclusion of “no evidence for MTX improving synovitis” appears explained by insufficient statistical power, wide individual variation, no subsets, low doses, and other limitations. MTX appears less efficacious in PsA than RA, but may be underestimated in PsA, similar to historical problems in RA, resulting more from limitations of clinical trials than from limitations of MTX.

The randomised, controlled clinical trial is appropriately regarded as the optimal method to establish the efficacy of a drug compared to a placebo or another drug (1). Many physicians and other health professionals apply the term “evidence-based medicine” almost exclusively to randomised controlled clinical trials and meta-analyses of these trials (2, 3). At the same time, important limitations are seen to the clinical trial methodology, as to any scientific method, particularly in chronic diseases over long periods, described in a number of reports (4-20), including previous commentaries by the senior author (21-29). However, most health professionals and the general public continue to ignore limitations of clinical trials. Increasing attention has been directed to limitations of clinical trials in recent years, including by experts in “evidence-based medicine” (16-20). A member of the Oxford Centre for evidence-based medicine commented that “while they are simple and easy to use, early hierarchies that placed randomised trials categorically above observational studies were criticised (17) for being simplistic (30). In some cases, observational studies give us the ‘best’ evidence. For example, there is a growing recognition that observational studies – even case-series (31) and anecdotes (32) can sometimes provide definitive evidence (17).”

Limitations of clinical trials may be seen in the convoluted history of methotrexate (MTX) for treatment of inflammatory rheumatic diseases since 1964 (33). Indeed, limitations of clinical trials, rather than limitations of MTX, delayed its introduction into management of RA 15–20 years after 1964. In PsA, MTX appears to have limited efficacy for axial manifestations (34), but possi-
ble benefits of MTX for synovial manifestations appear underestimated, again possibly due as much to limitations of clinical trials as to limitations of MTX. This article summarises some examples of limitations of clinical trials which may have affected development of MTX for RA or PsA over the last 50 years. All the studies cited were performed according to recognised standards for trial design and clinical care, and there is no intention to criticise any reports or authors cited. On the contrary, we hope these comments concerning MTX in clinical trials as well as observational data, may contribute to optimise therapies for patients with PsA, as well as other rheumatic diseases.

A. Limitations of clinical trials

Limitations of randomised control trials (Table I) include pragmatic and intrinsic limitations (27). Pragmatic limitations could be eliminated, in theory, by changes in design and implementation, while intrinsic limitations exist even in an ideally designed and executed clinical trial. Pragmatic limitations include: 1. A short time frame may limit an opportunity to recognise important trends in the long-term effectiveness, safety and tolerability of a medication. For example, in RA, gold and penicillamine were overestimated, and MTX underestimated in a meta-analysis of clinical trials (35), as most trials were of one year or less, and long-term loss of efficacy and adverse events could not be documented. Another example involved recognition of different outcomes of cyclophosphamide versus prednisone in nephritis of systemic lupus erythematosus (SLE), which required 5–10 years of observation, as no differences were seen after 1 year and even 3 years (36).

2. Incorrect and inflexible dosage schedules for test therapy as well as concomitant therapies; for example, in PsA, doses appear too high in the first methotrexate (MTX) trial (33), while doses appear too low in subsequent PsA clinical trials versus placebo (37-39).

3. Surrogate markers may not be optimal for long-term outcomes, due to over or under estimates of insufficient power, wide variation of individual patients.

Table I. Limitations of clinical trials in rheumatoid arthritis, other diseases, and psoriatic arthritis. MTX versus placebo trials.

<table>
<thead>
<tr>
<th>Pragmatic limitations</th>
<th>Rheumatoid arthritis (RA)</th>
<th>Other diseases</th>
<th>Psoriatic arthritis (PsA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relatively short observation period in which short-term efficacy may underestimate or overestimate long-term effectiveness</td>
<td>Meta-analysis that gold, penicamime, Ssz (35), trial of Mtx vs. auranofin (65) similar efficacy over one vs. 5 years</td>
<td>SLE nephritis trials (36) no effect at 1-4 years seen at 5 years</td>
<td>Most trials – weaker effects of MTX than in observational data</td>
</tr>
<tr>
<td>Incorrect, inflexible dosage schedules for test therapy as well as concomitant therapies</td>
<td>Too low for MTX</td>
<td>Too high (33) or often too low (37-39)</td>
<td></td>
</tr>
<tr>
<td>Surrogate markers often suboptimal for long-term outcomes</td>
<td>TJC and lab tests less significant than physical function to predict mortality (42) &amp; detect active biological vs control treatment (45)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistically significant may be clinically unimportant and vice versa, due to over or under estimates of insufficient power, wide variation of individual patients</td>
<td>Meta-analyses (35) and trial (65) on similar efficacy of DMARDs to MTX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insufficient power to recognise differences of MTX &amp; Pbo (33, 37-39)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variables other than randomisation may affect results</td>
<td>Education in BHAT study (110)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary analyses not on subtypes of PsA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intrinsic limitations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design of trial influences results – control group does not eliminate bias</td>
<td>Early RA (48, 78, 113) vs. long-term, failed MTX (114)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited data regarding subsets (33, 37-39)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data from clinical trials reported in groups – individual variation ignored</td>
<td>Almost all trials other than probability plots (79)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All trials – no probability plots (79)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balance of efficacy versus adverse effects not standard</td>
<td>Trials of mammography in women age 40-50 (52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All trials</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Placebo effect” distorted</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All trials</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Limitations of MTX clinical trials / T. Pincus et al.
counts or laboratory tests (45), but may be dismissed as “subjective.”
4. Statistically significant results may be clinically unimportant and vice versa; many results in PsA trials appear to present clinically important advantages to MTX versus placebo (37-39), although they do not meet the $p<0.05$ criterion, in part due to insufficient statistical power and non-stratification of patients regarding the proportion of patients with polyclinical versus axial disease.

5. Variables other than randomisation may affect outcome; for example, in the beta blocker to prevent second heart attack trial (BHAT) in patients who had suffered a myocardial infarction, formal education level of the patient was far more significant to predict mortality outcomes than whether the patients were randomised to a beta blocker or placebo (46).

6. Inclusion and exclusion criteria often result in eligibility of only a small minority of patients with a disease; for example, fewer than 10% of people with RA in the author’s care before 2000 were eligible to participate in most trials of biological agents (47).

Intrinsic methodologic limitations of clinical trials are seen even if the trial could be designed and executed ideally:

1. The design of the clinical trial can greatly influence results; a control group does not necessarily eliminate bias. For example, a positive treatment effect to MTX is more likely in patients who had no previous MTX such as in the PREMIER trial (48) compared to patients who “failed” two previous treatments such as in the ARMADA trial (49).

2. Data from a clinical trial are reported in groups, and results do not necessarily pertain to all individual patients; for example, in trials to compare acetaminophen to non-steroidal anti-inflamma
dory drugs (NSAIDs) in osteoarthritis (OA) (50, 51), 55-60% of patients preferred diclofenac/misoprostol, but 20-25% of patients preferred acetaminophen (20-25% reported similar ef
cfficacy to both agents) (50). Results in some individuals differ from those in the majority of patients in the majority of clinical trials.

3. A clinical trial cannot definitively estimate adverse events, which may lead to different interpretations of data; for example, clinical trials in mammography in 40-50 year old women can give different interpretations of results, depending on assumptions, say, about the morbidity and mortality of “false positive” biopsies versus missed lesions, etc. (52).

4. The placebo effect of treatment is distorted, and may be exaggerated or attenuated when a patient is told that she/he is receiving a therapy in a “scientific experiment,” versus the “best therapy” from her/his doctor (14).

B. Underestimation of methotrexate in rheumatology due to limitations of clinical trials: 1964-1985

The first clinical trial of MTX in any rheumatic disease was conducted in patients with PsA and reported in 1964 (33). In this trial, 21 PsA patients were given a course of 3 intravenous MTX injections at 10 day intervals in doses of 1, 2, and 3 mg/kg or 3 placebo injections in a crossover design. All 21 patients had involvement of small joints of the hands, all but one of distal interphalangeal (DIP) joints; 12 had spine involvement, most frequently the cervical spine, and 15 had knee involvement. All but one patient was seronegative for rheumatoid factor, and all but one patient had received systemic glucocorticoid therapy (33).

Significant advantages to MTX versus placebo were seen for joint involvement, skin involvement erythrocyte sedimentation rate (ESR), and joint range of motion ($p<0.01$) (33). The authors noted that “the clinical response in most patients … was quite impressive … many totally disabled patients were restored to an ambulatory and employable state (33).”

However, adverse events were common, including anorexia in 14 of 21 patients, nausea in 13, burning sensa
tion in the skin in 10, leucopenia in 7, anaemia in 3, transient elevation of transaminase (SGOT) in 3, and thrombocytopenia in 1. One of the 21 patients died, who had been enrolled in the study after having received prednisonae 30 mg/day for 8 months (33). The pa
tient was then treated with intravenous doses of 50, 100, and 150 mg MTX at 10-day intervals. After the third in
dication, the patient’s white blood cell count fell to 400/cu mm, and platelet count to 18,000/ cu mm. On the 13th post-injection day, he developed hae
morhages and died. Autopsy revealed oesophagitis and a pulmonary embolus.

Although his bone marrow had recovered fully, his death was attributed to the MTX treatment (33).

The results of the clinical trial were inter
terpreted as indicating efficacy of MTX in psoriasis and PsA, but with many adverse events “suggesting unaccepta
ble toxicity” (39). In retrospect, toxicity appears explained by higher doses of MTX and prednisone 2 to 5 times more than the maximum that would be administered at this time. The patients were not treated with folic acid, as the action of MTX at the time was thought secondary to anti-folate, anti-
metabolite properties, rather than to an
ti-inflammatory activity, as recognised today (53, 54).

One response to the results of the 1964 trial might have been an effort to de
velop a safer regimen of methotrexate, in view of documented significant effic
acy (33). However, MTX was largely abandoned by the rheumatology com
munity over the next 15 to 20 years.

MTX was regarded by almost all rheu
matologists as a “cancer drug” that was too toxic and unnecessary for RA, while injectable gold salts and peni
cillamine were considered “remission
inducing therapy” as recently as 1985 (55). Ironically, designation of gold salts and penicillamine as “remission-inducing” and rare use of MTX from 1964 to 1979 occurred largely on the basis of limitations in design and inter
pretation of clinical trials (28, 56), and not on the basis of limitation of MTX.

A few pioneering rheumatologists such as Hoffmanstein (57, 58), Scherbel (59), and Zachariae (60) treated patients who had inflammatory rheumatic disease with MTX during the 1970s and 1980s. An abstract presented at the American Rheumatism Association [ARA - now American College of Rheumatology (ACR)] by Hoffmanstein in 1972 stated “Twenty-nine patients with classic or definite adult rheumatoid arthritis (RA) have been treated with a single 10 to 15 mg dose of methotrexate
once every 7 days. Most were on 5 mg prednisone daily, or less. Average duration of methotrexate therapy was 25.5 months... Improvement of functional capacity was moderate in 7, mild in 15 and absent in 7. Clinical improvement was estimated as major in 11, moderate in 14, and minor or none in 4 (57).” These results 43 years ago appear quite similar to those in contemporary care. Hoffmeister’s observations were not welcomed by the leadership of the rheumatology community (personal communication), in large part as the data were not from a clinical trial. Furthermore, as noted, gold salts and penicillamine were regarded as “remission inducing,” largely as a result of interpretation of clinical trials conducted over one year, in which some patients had disappearance of signs of RA (55). However, over the 1970s and 1980s, it was increasingly recognised that gold salts and penicillamine were associated with substantial loss of efficacy over several years, as well as with rare, but serious and sometimes fatal, long-term haematologic, renal, and bone marrow adverse events. Remission over longer than 3 years was seen in fewer than 2% of patients treated with these agents (61), and gold salts and penicillamine are rarely used at this time.

Underestimation of MTX in rheumatology due to limitations of clinical trials: 1985-2000

Two clinical trials were conducted in patients with RA in the 1980s indicating the efficacy and safety of weekly low-dose MTX versus placebo (62, 63). These trials followed (rather than preceded) observational data from routine clinical care (57-60) in rheumatology and dermatology (60). For example, MTX had been introduced to treat 35% of patients with RA under care of the senior author by 1985 (Fig. 1) (64). Despite evidence of efficacy from these two clinical trials (62, 63) and observational data (57-60), doubts concerning MTX compared to “remission-inducing” DMARDs continued to be reported to the rheumatology community. For example, a meta-analysis published in 1990 (Fig. 2) analysed the efficacy and safety in 66 clinical trials of DMARDs that included 117 treatment groups (35). The composite treatment effect of MTX (7 trials) was indistinguishable from sulphasalizine (6 trials), d-penicillamine (19 trials), and parenteral gold (29 trials); antimalarials (11 trials) had a lower treatment effect, with still lower levels for auranofin (oral gold), and lowest for placebo (22 trials). Dropout rates were significantly higher for parenteral gold than all other DMARDs, and the lowest dropout rates (lower than placebo) were seen for anti-malarials and MTX (Fig. 2) (33). Two-thirds of the 66 trials in the meta-analysis were reported during the 1980s, including all 7 involving MTX, indicating that most had been designed after Hoffmeister’s 1973 abstract (35).

In 1992, two years after the meta-analysis report, a multicentre clinical trial was conducted to compare MTX, auranofin (oral gold), or the combination of the two drugs in 335 patients over 48...
weeks (65). No significant differences were seen in any clinical or laboratory variable (Fig. 3) (65). The authors tempered their observation of no advantages to either single drug or a combination with a comment “within the time frame of the study (65).” The conclusions of the meta-analysis and clinical trial in Fig. 2-3 appeared inconsistent with observational data from 7 practices in clinical care, also reported in 1992, concerning continuation of courses of DMARDs in RA patients over 5 years (Fig. 4) (66). Continuation of a therapy may best represent a decision that the short-term efficacy and long-term effectiveness of a therapy justifies issues of tolerability and potential toxicity in analysis of benefit/risk by doctor and patient. More than 50% of courses of MTX were continued at 60 months (5 years), compared to 20% of courses of hydroxychloroquine, penicillamine, parenteral gold and azathioprine and fewer than 10% of courses of auranofin or oral gold (66) (Fig. 4a) (p<0.01). However, a sub-analysis over one year (rather than 5 years) of only the patients who had prior DMARDs (to simulate clinical trials) indicated no clinically or statistically significant differences between MTX, hydroxychloroquine, penicillamine, parental gold, azathioprine, or oral gold (Fig. 4b).

These analyses indicated that results over one year in clinical trials, including the extensive meta-analysis, are also seen in actual care, but the results are not applicable to data over 5 years in clinical care. Further observational studies indicated that continuation of courses of weekly low-dose MTX is almost always longer than seen for any therapy for rheumatic diseases, other than possibly prednisone (67), including biological agents (68). Perhaps the most instructive example of the need for long-term observations in rheumatology clinical trials is the “exception that proves the rule,” a trial of cyclophosphamide (Fig. 5) (36). This trial is one of the few rheumatology clinical trials conducted over as long as 3 years, and indicates the need for long term observations to recognize certain treatment effects in chronic diseases. Cyclophosphamide might not have been established as the standard of care over the next 2 or 3 decades for the treatment of SLE nephritis if the trial had been terminated after one or even 3 years.

The problem of short time frame may be relevant to many rheumatology clinical trials over the years, including at this time, which appear to indicate negative or equivocal results. Of course, many therapies that do not differ from placebo or another therapy after one year also will not differ over longer periods. Nonetheless, Figure 5 might be recalled when similar results are seen over one year for a therapy versus another therapy or placebo in a rheumatic disease, including PsA (33, 37-39, 69, 70), as well as SLE (71, 72), systemic sclerosis (73-75), and other rheumatic diseases. These trials often are regarded as “failed,” but interpretation of results may be complex. This concern may be particularly relevant when many measures indicate advantages to a therapy which is not statistically significant according to the p<0.05 criterion, as seen in 3 of 4 PsA trials (37-39) [other than the first trial (33) in which difference between MTX and placebo were significant, but the dose was far too high]. MTX became the standard of care for RA and emerged as the “anchor drug” for RA over the last decade (76, 77), and is recommended for PsA (34), in part because of clinical trials, but largely by overcoming limitations of clinical trials and interpretations of results.

**Estimated Continuation of Courses of 2nd-Line Therapy**

- **a) All Courses Over 60 Months**
- **b) Initial Course Over 12 Months**

**Fig. 4.** Estimated continuation of courses of 6 DMARDs in 477 patients with rheumatoid arthritis in 7 rheumatology practices (4): a) 532 courses over 5 years; b) 477 initial courses over 1 year (66).
Of course, it is not possible to maintain randomisation over 5–10 years in symptomatic patients with RA, PsA, or any chronic disease, for ethical reasons, not to mention costs. However, open-label extensions for indefinite periods to analyse outcomes such as renal failure, joint replacement, work disability or death are possible at low costs using, the internet, telephone, or mail, without a need for costly patient visits. Consent for and implementation of such long-term extensions might be required in clinical trials of all chronic diseases.

C. Underestimation of MTX in rheumatology due to limitations of clinical trials: 2000-2010

In the early 2000s, several clinical trials were conducted by pharmaceutical companies to compare radiographic progression in patients treated with biologic agents or MTX or a combination of MTX and the biological agent. For example, the TEMPO trial of early RA patients indicated radiographic progression of 3.34 units over 2 years in patients randomised to MTX, compared to 1.15 in patients randomised to etanercept, and 0.56 in patients randomised to the combination of MTX and etanercept (78). These differences are statistically significant. However, a change of 3.34 units is less than 0.5% of the total number of units in a Sharp/van der Heijde score of 448 (79) (Fig. 6). Patients rarely develop more than 50% of the maximum score and it may be suggested that maximum scores actually are 224, or 50% of maximum. Nonetheless, 3.4 units over 2 years would not be detectable in an individual patient (80), and a clinically detectable change of 17 units (5 x 3.4) would be seen only after 10 years would not be useful in actual patient care.

Furthermore, hand radiographs are limited as predictors of severe RA outcomes such as work disability (40, 41) or mortality (42). A review of all 53 reports which described long-term predictors of premature mortality in RA indicated that functional status was significant in analyses in 17 of 18 studies, while hand radiographs were significant in 5 of 18 studies (42) (Fig. 7). Although structural damage is correlated significantly with functional disability, the levels of r=0.3–0.5 indicate that radiographs explain less than 25% of variation in functional disability (81). Results similar to TEMPO are seen for many biologic agents (Fig. 6) (56), with statistically significant differences between the biological agent and MTX in early RA patients, which would be clinically undetectable in individual patients (56). Moreover, MTX inhibits radiographic progression considerably (82) and remains the mainstay therapy for RA, taken by far more patients than any DMARD or biological therapy, despite observations of radiographic differences versus biological agents (83).
The observation in MIPA that 2 of the 7 ACR Core Data Set measures indicated statistically significant advantages to MTX over placebo, i.e., physician global estimates and patient global estimate. This finding was interpreted as “symptom-modifying,” as the more “objective” measures of laboratory tests and joint counts, as well as indices, were not statistically significant in differences between MTX and placebo (39). However, physician and patient global estimates are most likely among the 7 RA Core Data Set to discriminate active from control treatment in clinical trials of biological agents, which document disease-modifying activity, defined as radiographic progression (45). Therefore, differences according to the global measures in the MIPA study may have some clinical significance (in contrast to statistical significance), may be more than acknowledged.

Analyses of data from all 9 clinical trials in which relative efficacies of core data set measures to distinguish active from control treatments have been computed (of MTX, leflunomide, adalimumab, abatacept, and infliximab) indicates that physician global estimate and patient global estimate are the most efficient measures (Fig. 8) (45). Analyses of 4 studies involving adalimumab, ARMADA, DE011, STAR, and DE019 (85), indicated that physician global estimate had the highest relative efficiency to distinguish active from control treatments (45). In 3 studies of MTX or leflunomide versus placebo (86, 87), and infliximab versus MTX (45, 88), patient global estimate had the highest relative efficiency to distinguish active from control treatments (45).

The measures among 7 RA Core Data Set measures that were among the 3 highest in relative efficiencies in the 9 studies for which data have been analysed, which might be chosen for an optimal index to assess improvement in clinical trials (89), were physician global estimate in 7/9 (78%), patient global estimate, pain, and HAQ function in 5/9 (56%), SJC and ESR/CRP in 3/9 (33%), and TJC in none (Fig. 8A) (45). Analyses only of 3 measures, one from each category of the core data set – laboratory, assessor, and patient,
included the single laboratory test – ESR or CRP, whichever had the higher efficiency if both were available, SJC—the joint count measure with higher relative efficiency (vs. TJC), and HAQ physical function—which is regarded as the least likely among RA Core Data Set measures to be reversible (90, 91).

In these analyses, HAQ physical function was most efficient (Fig. 8B) (45), although each was significant in more than one of 9 trials, indicating the need for an index. An emphasis on radiographic scores and joint counts over functional disability as a primary outcome measure in clinical trials reflects dominance of a “biomedical model” (92, 93) in contemporary medicine. A major feature of this model is that “objective” data from high technology sources or detailed examinations are more important than data from a patient. The importance of a “biomedical model” is seen in dismissal of statistically significant differences between MTX and placebo in patient and physician global assessment as “subjective” in the MIPA trial (39).

Another example of the preeminence of a biomedical model in rheumatology may be seen in a requirement for 20%, 50%, and 70% improvement in TJC and SJC versus only 3 of the other 5 RA core data set measures to meet ACR 20, 50, and 70 responses (43). Therefore, TJC and SJC are weighted as “objective” measures higher than a physician global estimate or patient self-report scores. More recently, an ACR/EULAR committee to establish new criteria for remission in early RA made an a priori decision that any criteria for remission must include a TJC, SJC, and laboratory tests (94). However, no “evidence” is available that favourable values for joint count or laboratory tests are more prognostic of long-term outcomes such as work disability or premature death, or more likely to distinguish active from control treatment more effectively than the other 4 RA core data set measures, i.e. 3 patient self-report measures of physical function, pain, patent global assessment and physician global assessment in clinical trials. The decisions concerning ACR improvement and remission criteria are not “evidence-based” from the standpoint of long-term outcomes or clinical trial results. However, these decisions may be appropriate on the basis of clinical considerations.

To their credit, the authors of the MIPA trial present several supplementary tables, as well as a candid description of possible limitations of their study. An analysis of completers only (rather than all participants in “intention to treat” analyses, as was presented in the report) indicated a statistically significant advantage to MTX according to the PsARC primary outcome (95), as well as higher odds ratios according to the ACR 20 and DAS28 secondary outcomes (39). However, the authors point out that the intent to treat analyses in the published report are preferable methodologically, which is technically correct, but completer analyses may be preferable in some instances (perhaps in MIPA) because of issues of statistical power, short time-frame, and absence of PsA subset analyses. Indeed, another supplementary table classifies participants into two cate-
lies, oligoarticular (79 patients) and polyarticular (142 patients), including a further subset of polyarticular termed "RA equivalent" (37 patients). The differences between MTX and placebo-treated patients were considerably greater in the polyarticular, than oligoarticular patient group and dramatically different in the "RA-equivalent" group (39). The authors commented that "interpretation of these data is fraught with difficulty and requires extreme caution as the analyses were not pre-determined, the groups were very small, the confidence intervals are large. Nevertheless, we have included the data online for interest" (39). Perhaps, it would have been desirable that these analyses would have been included in the main report, rather than as a supplementary table; perhaps they were included in the initial submission, and changes were made during the editorial process.

The authors also discussed possible limitations of MIPA, suggesting that "it is possible that MTX might be effective in some circumstances in patients with PsA" (39). The first involved possible higher doses of MTX for longer periods, which was not thought important as "placebo-controlled RCTs show 15 mg weekly or less MTX is effective within three months" (39). The second was that inclusion criteria allowed patients with fewer active joints than other PsA trials, but this was not regarded as a problem since “most patients were more active than the minimum requirement and MIPA was designed as a pragmatic trial” (39). Thirdly, there is a concern regarding sample size, which is explained by other trials having higher odds ratios using the imputation methodology in this study. The MIPA authors suggest that studies with apparently conflicting results concerning MTX efficacy may have included a greater proportion of patients with RA-like PsA (39).

The report of the MIPA trial notes five previous clinical trials involving MTX in PsA, 3 versus placebo and 2 versus cyclosporine, as supporting the absence of MTX efficacy in PsA. However, each of these 5 trials showed meaningful efficacy of MTX. The first trial (33) as discussed above, was noted to have “unacceptable toxicity”, but no comment was made that the doses were 2–5 times more the maximum at this time of 30 mg/wk. A second trial included 37 patients (37); as in MIPA, physician global assessment and skin psoriasis were improved at a significantly statistical level, and all measures other than TJC showed an advantage to MTX. A third trial of 35 patients with early PsA, all of whom were described as “belonging to the oligoarticular subset” showed statistically significant advantages to MTX for SJC, TJC, patient and physician global estimates, and pain visual analogue scale (38). The other two clinical trials documented similar improvement with cyclosporine and MTX (70). The MIPA discussion also commented that systematic reviews also are consistent with a “paucity of evidence” concerning possible efficacy of MTX in PsA (96, 97).

Several rheumatologists also questioned some interpretations of findings in MIPA. An editorial in the same issue of Annals of Rheumatic Diseases pointed out that “such negative results are not in line with the physicians’ views”, and points to problems with inclusion criteria, the choice of the primary outcome and study duration (98). The editorial appears to accept the idea that MIPA is the "third negative placebo-controlled trial of MTX versus placebo," but suggested that since the number of involved joints was low, a more informative outcome measure might have been the number of patients with no swollen joints.

A letter to the editor in a subsequent issue (99) also raised issues with the inclusion criteria, suggesting that the CASPAR criteria (100) might have led to a different outcome. Several reviews specifically pointed out issues with MIPA, including a channeling bias for milder patients, low MTX dose, long period to dose escalation, insufficient statistical power to recognise differences between MTX and placebo, and high dropout rate (101-103). An article in an earlier Supplement in this series concerning MTX concluded that “it seems that MTX does have an important role in the management of PsA” (97).

The MIPA authors responded that the number of patients with no joint swelling was similar in both groups and that the same result would have been seen using the CASPAR criteria (104). They also suggested that physicians’ acceptance of MTX as beneficial may be explained by “regression to the mean” (104).

An open-label study of 115 PsA patients who had not received MTX found ACR20 responses in 67% receiving MTX only compared to 86% of patients who received infliximab plus MTX (105). Treatment was more efficacious in the infliximab group (as expected); however, 67% ACR 20 responses is in the range seen in the treatment arm of RA clinical trials. It may be noteworthy that the initial SJC and TJC, as well as ESR and CRP, were higher than in the MIPA trial (39, 105).

In the tight control of inflammation in early PsA trial (TICOPA) (106), advantages were seen to the tight control group. MTX monotherapy was continued in 27% patients in tight control group and 60% in the standard care group. No differences were seen in radiographic progression, which was quite low in both groups, suggesting that MTX is “disease-modifying” in PsA as in RA. The TICOPA authors noted that “MTX is often the first drug to be used in PsA at doses that are considerably higher than in the MIPA trial, with 82% in the tight control group taking 25 mg weekly by 12 weeks” (106). A study from Norway reported in 2009, indicated comparable improvements in 330 patients with PsA compared to 929 with RA, with no significance difference in SJC, ESR, CRP, or Physician global VAS, and somewhat greater in TJC (3.9 vs. 3.2) and DAS 28 (1.38 vs. 1.13). Drug retention was similar in both groups at two years with 65% and 66% continuing MTX therapy in PsA and RA respectively (107).

Concluding comments

The randomised controlled clinical trial remains the most effective approach to distinguish an active treatment from a placebo or other treatment. However, clinical trials have a number of limitations as do all scientific methods. Further refinement of clinical trials over the last few decades has been focused more on statistical methodology than on clinical relevance, and the possibil-
ity that results may be misleading despite a control group and excellent statistical methods.

Possible advantages of a treatment versus a placebo which does not meet the p<0.05 criterion, as seen in most trials of MTX in PsA (37-39), may not necessarily indicate the absence of a treatment effect. The findings may result from limitations of the clinical trials, including inadequate statistical power, selection of patients with mild disease, low doses and slow escalation of MTX, and the absence of pre-specified stratification of patients regarding primarily axial, oligoarticular, symmetrical, or polyarticular presentation.

A negative interpretation of MTX efficacy in a clinical trial in 1964 (33) was a major contributor to 15-20 years in which patients with RA, PsA and other rheumatic diseases were not routinely treated with MTX, the current standard of care for RA. In PsA, MTX appears less effective than RA, as axial disease usually does not respond. Nonetheless, the efficacy of MTX in PsA may be underrated in polyarticular arthritis, although polyarticular arthritis in PsA differs from RA (108).

Observational data may provide insights not available from clinical trials, particularly when all consecutive patients seen are captured in the data set (109), which may be enhanced with electronic data capture for patients. Data from patients appear as informative as data from physical examination, imaging, and laboratory to distinguish active from control treatments in RA and to predict severe long-term RA outcomes such as work disability and mortality.

Although many rheumatologists talk about treatment of most diseases often with a suggestion of a need for (more) clinical trials, most problems in rheumatic diseases have not been and will not be solved through clinical trials, due to costs and other pragmatic limitations. The rheumatology community has ethical and intellectual responsibilities to attempt optimal capture of data from routine care to improve therapies toward better outcomes for our patients with all rheumatic diseases, most of which will not be derived from clinical trials.

References
29. PINCUS T, CASTREJON I: Evidence that the strategy is more important than the agent to treat rheumatoid arthritis - data from clinical trials of combinations of non-biologic DMARDs, with protocol-driven intensification of therapy for tight control or treat-to-target. Bull Hosp Jt Dis (2013). 2013; 71 (Suppl. 1): 33-40.
34. GOSSEC L, SMOLLEN JS, GAUFOUX-VIALA C et al.: European League Against Rheumatism recommendations for the management of psoriatic arthritis with pharmacological
Limitations of MTX clinical trials / T. Pincus et al.

35. FELSON DT, ANDERSON JJ, MEENAN RF: The comparative efficacy and toxicity of sec-
ond-line drugs in rheumatoid arthritis. Re-
37. WILLKENS RF, WILLIAMS HJ, WARD JR et al.: Randomized, double-blind, placebo con-
38. SCARPA R, PELUSO R, ATTENO M et al.: The effectiveness of a traditional therapeu-
43. FELSON DT, ANDERSON JJ, BOERS M et al.: American College of Rheumatology. Pre-
liminary definition of improvement in rheu-
44. FELSON DT, ANDERSON JJ, BOERS M et al.: The American College of Rheumatolog-
45. PINCUS T, RICHARDSON B, STRAND V, BERGMAN MJ: Relative efficiencies of the 7 rheumatoid arthritis Core Data Set measures to distinguish active from control treat-
46. RUBERMAN W, WEINBLATT E, GOLDBERG JD, CHAUDHARY BS: Psychosocial influ-
ces on mortality after myocardial infarc-
47. SOKKA T, PINCUS T: Most patients receiv-
ing routine care for rheumatoid arthritis in 2001 did not meet inclusion criteria for most recent clinical trials or American Col-
48. BREEDVELD FC, WEISMAN MH, KAVAN-
AUGH AF et al.: The PREMIER study: A multicenter, randomized, double-blind clini-
cal trial of combination therapy with adali-
 mumab plus methotrexate versus methotrex-
ate alone or adalimumab alone in patients


