Relative efficiencies of the 7 rheumatoid arthritis Core Data Set measures to distinguish active from control treatments in 9 comparisons from clinical trials of 5 agents

T. Pincus[1], B. Richardson[2], V. Strand[3], M.J. Bergman[4]

[1]Rush University School of Medicine, Chicago, IL, USA;
[2]University of Sydney, Sydney, NSW, Australia;
[3]Stanford University, Stanford, CA, USA;
[4]Drexel University, Philadelphia, PA, USA.

Theodore Pincus, MD,
Benjamin Richardson
Vibeke Strand, MD
Martin J. Bergman, MD

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

Received on October 10, 2014; accepted in revised form on October 11, 2014.

Clin Exp Rheumatol 2014; 32 (Suppl. 85): S47-S54.

Key words: rheumatoid arthritis, Core Data Set measures, clinical trials, patient questionnaires, joint counts

Competing interests: none declared.

ABSTRACT

The 7 Core Data Set measures to assess rheumatoid arthritis (RA) were analysed for their relative efficiencies to distinguish active from control treatments in 9 comparisons of 5 agents, methotrexate, leflunomide, infliximab, adalimumab, and abatacept, in 8 clinical trials. Among the 7 measures, levels of relative efficiencies were in a similar range, highest for the physician global estimate, followed by, in order, patient global estimate, physical function on a health assessment questionnaire (HAQ), pain, swollen joint count (SJC), an acute phase reactant laboratory test – erythrocyte sedimentation (ESR) or C-reactive protein (CRP), and tender joint count (TJC). Comparisons of only 3 measures, SJC and ESR/CRP (regarded as optimal indicators of inflammation) and HAQ function (regarded as most likely to be affected by joint damage and therefore least reversible) indicated relative efficiencies for HAQ function at least as great as for SJC or ESR/CRP, although 8 of the nine comparisons involved patients with disease duration >6.9 years. The findings indicate a strong rationale for a Core Data Set of 7 measures, as no single measure was clearly superior in relative efficiency in all clinical trials. At the same time, "objective" laboratory ESR/CRP, TJC and SJC were not superior to "subjective" global estimates of the physician or patient or patient self-report measures of physical function or pain, to differentiate active from control treatments. The findings challenge a traditional view that laboratory and clinical examination findings are more robust than patient self-report scores and physician global estimates to assess and monitor RA patients.

The Core Data Set for assessment of rheumatoid arthritis (RA) (1) includes 7 measures: 3 from the physical examination - swollen joint count (SJC), tender joint count (TJC), and physician estimate of global status (DOCGL); 3 from the patient history information –physical function (FN), pain (PN), and patient estimate of global status (PATGL); and one laboratory test – erythrocyte sedimentation (ESR) or C-reactive protein (CRP). A radiographic score is added when a study involves more than one year (and more recently 6 months) of observation (1).

The composition of the Core Data Set reflects the prominence of the patient history and physical examination, in contrast to vital signs, laboratory tests and ancillary studies, in clinical decisions in RA compared to many other chronic diseases (2). A survey of 313 physicians, including 159 non-rheumatologists and 154 rheumatologists (who did not differ substantially in their estimates) (2) indicated that RA was the only one of 8 chronic diseases for which the patient history and physical examination each provided more than 50% of the relevant information for diagnosis and management (total >100% due to "ties" in estimates of proportions of relevant information by physicians). By contrast, clinical decisions in many other diseases are dominated by a *gold standard* biomarker, such as blood pressure in hypertension, haemoglobin A1C in diabetes, or a biopsy in lymphoma (2, 3).

The 7 RA Core Data Set measures were not designed to be weighted for greater or lesser importance of any individual measure. However, in many rheumatology indices and criteria, SJC, TJC, and ESR or CRP are regarded as more important "objective" measures than patient reported "subjective" physical function, pain, or global estimates by physicians or patients, reflecting the preeminence of a *biomedical model* (4). For example, a 20%, 50%, or 70% response according to American Col-

Table I. Relative efficiencies to distinguish active from control treatments (rank among 7*) of 7 Core Data Set measures in a comparison in clinical trials.

Comparison	Leflunomide *vs*. Placebo	MTX *vs*. Placebo	Adalumimab *vs*. MTX	Adalumimab *vs*. Placebo	Adalumimab *vs*. DMARD	Adalumimab *vs*. MTX	Infliximab *vs*. MTX	Infliximab *vs*. MTX	Abatacpt *vs*. MTX
Reference of Trial	Strand 1999 (11)	Strand 1999 (11)	Weinblatt 2003 (14)	Van de Putte 2003 (16)	Furst 2003 (17)	Keystone 2004 (15)	Maini 1999 (12)	St. Clair 2004 (13)	Genovese 2006 (18)
Reference of Relative Efficiencies	Tugwell 2000 (7)	Tugwell 2000 (7)	Pincus 2008 (9)	Pincus 2008 (9)	Pincus 2008 (9)	Pincus 2008 (9)	Pincus 2009 (8)	Pincus 2009 (8)	Wells 2008 (10)
	US301	US301	ARMADA	DE011	STAR	DE019	ATTRACT	ASPIRE	ATTAIN
Median age (years)	54.1-54.6	53.3-54.6	53.5-58.2	50.2-53.7	55.0-55.8	56.1-57.3	51-56	50-51	52.7-53.4
Median duration of disease (years-rank of 9 comparisons)*	6.9-7.0 (6)	6.5-6.9 (8)	11.1-13.1 (1)	9.4-10.4 (5)	9.3-11.5 (4)	10.9-11.0 (3)	7.2-9.0 (6)	0.8-0.9 (9)	11.4-12.2 (2) (2)
TJC	1 (5)	1 (5)	1 (5)	1 (7)	1(6)	1 (6)	1 (6)	1 (7)	1 (6)
SJC	0.56 (7)	0.91 (6-7)	**2.2 (2)**	1.55 **(5)**	1.12 **(5)**	1.42 **(4)**	**1.63 (3)**	**2.79 (1)**	0.57 (7)
DOC GL	1.33 (3)	**1.44 (2)**	**2.72 (1)**	**2.65 (1)**	**1.66 (1)**	**2.06 (1)**	1.48 **(4)**	**2.54 (3)**	1.16 **(4)**
ESR	0.48(NR)	1(NR)	NA	NA	NA	NA	NA	NA	**1.82 (1)**
CRP	0.63 **(6)**	**1.19 (3-4)**	**1.86** (3)	1.3 **(6)**	0.22 **(7)**	0.6 (7)	0.41 (7)	1.91 **(4)**	0.93 **(NR)**
HAQ FN	**1.84 (2)**	0.91 (6-7)	0.94 **(6)**	1.6 **(4)**	**1.27 (3)**	**1.52 (2)**	1.41 **(5)**	**2.6 (2)**	**1.22 (3)**
PAIN	1.21 **(4)**	**1.19** (3-4)	0.92 **(7)**	**2.12 (3)**	1.17 **(4)**	**1.48** (3)	**2.78 (2)**	1.41 **(6)**	**1.38 (2)**
PAT GL	**1.88 (1)**	**1.55 (1)**	1.48 **(4)**	**2.14 (2)**	**1.43 (2)**	1.36 **(5)**	**3.28 (1)**	1.74 **(5)**	1.04 **(5)**

MTX: Methotrexate; DMARD: disease modifying anti-rheumatic drug; TJC: tender joint count; SJC: swollen joint count; DOCGL: physician global estimate; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein; HAQ: Health Assessment Questionnaire Physical Function score; PATGL: patient global estimate; NR: not ranked; NA: not available. *Ranked highest to lowest.

lege of Rheumatology (ACR) criteria requires at least that degree of improvement in SJC and TJC, along with similar improvement in 3 of the other 5 measures (5); a patient global estimate may be improved from say, 8/10 to 1/10, but if the SJC is unchanged, the patient is regarded as not meeting ACR20 criteria. ACR/EULAR (European League against Rheumatism) remission criteria for RA based on the Core Data Set require TJC≤1 and SJC≤1, and normal ESR or CRP, which were selected as 3 required measures before a search for an additional measure added patient global estimate ≤1 (6); a patient may have a pain score of 6/10, but be regarded as in remission.

Relatively little data are available to compare how informative each of the 7 Core Data Set measures might be to guide clinical care. One approach is to compare the relative efficiency of each measure to distinguish active from control treatments in randomised controlled clinical trials (7). Data are available concerning 9 comparisons of the relative efficiencies of the 7 Core Data Set measures in 8 clinical trials of 5 agents, methotrexate (7), leflunomide (7), infliximab (8), adalimumab (9), and abatacept (10). All relative efficiencies were calculated according to the method of Tugwell, Wells, Strand, *et al*. (7), in which the other 6 RA Core Data Set measures are compared to TJC as the referent measure (=1).

It would be ideal to analyse the 9 comparisons of relative efficiencies according to a meta-analysis, but that would require the original data from disparate sources, which are not available to the authors. Therefore the analyses were performed on the basis of a simple compilation of the reported relative efficiencies to distinguish active from control treatments (Table I). This report presents all 9 reported comparisons of relative efficiencies known to the authors from published clinical trials involving methotrexate (11), leflunomide (11), infliximab (12, 13), adalimumab (14-17), and abatacept (18) (Table I).

Methods

All 9 comparisons of the 7 Core Data Set measures were analysed for relative efficiencies in post hoc retrospective analyses according to the method of Tugwell, Wells, Strand, *et al*. (7). Initially, the standard effect size was calculated as a ratio in which the difference between the mean values of each active treatment *versus* control arms was the numerator, and the corresponding calculated standard deviation from analysis of variance for the two treatment groups was the denominator.Relative efficiency was then obtained by dividing the square of the standard effect size of the variable by its counterpart for the tender joint count as the referent measure.

In analyses of adalimumab data, in addition to using the published method (7) in which arithmetic changes for each Core Data Set measure were computed, percentage changes also were computed (9). However, in order to compare these 4 trials to the other trials, in which only the arithmetic means were available, only the arithmetic means from the adalimumab trials were incorporated into the rankings in this report. Some reports also included data from other measures than those in the RA Core Data Set (10, 19), such as data from the short form 36 (SF-36) (20), but these data are not included in the analyses performed in this report.

Relative efficiencies compared to TJC were compiled into a table (Table I), and ranked according to highest (=1) to lowest (=7) values in each of the 9 comparisons. If both ESR and CRP were available, the measure with the higher relative efficiency was regarded as the rank for the "laboratory variable." The ranks were then compiled into a composite score, based on whether the

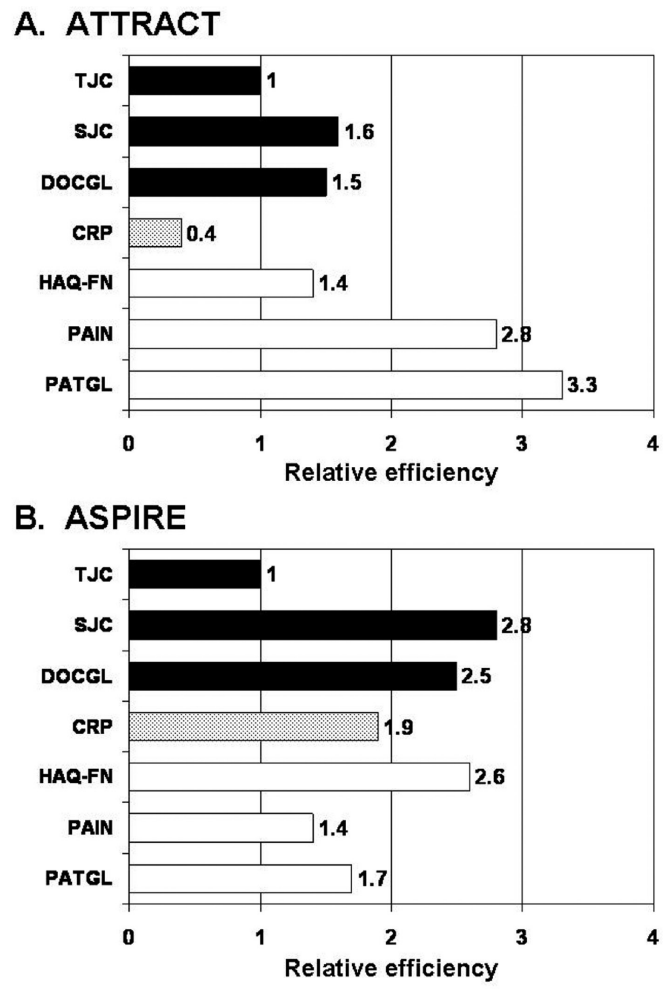

Fig. 1. Relative efficiencies of 7 Core Data Set Measures to distinguish patients with Infliximab versus control therapies in two clinical trials, ATTRACT (12) and ASPIRE (13). Relative efficiencies were compared to tender joint count (TJC) as reference measure.

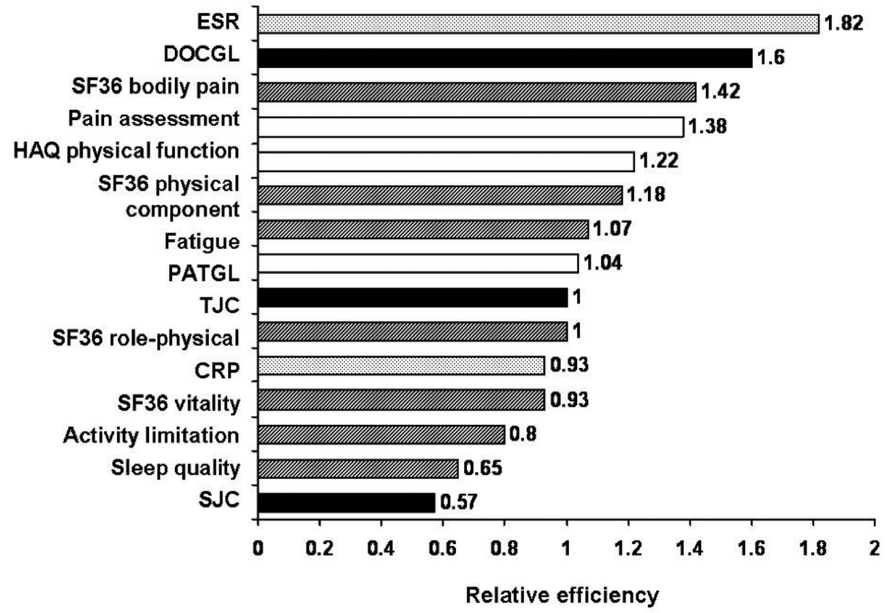

Fig. 2. Relative efficiencies of various measures to distinguish abatacept from control treatment in the ATTAIN clinical trial for RA. Relative efficiencies were compared to referent tender joint count. The figure contains the 7 Core Data Set Measures as well as scales from the Short Form 36 (20) as well as fatigue and sleep quality (10) which are not included in Table I. Only the 7 Core Data Set measures are included in Table I.

measure was among the highest 5, 4, 3, or 2 measures among the 7 Core Data Set measures in each of the 9 comparisons. For example, in analyses of relative efficiencies of leflunomide *versus* placebo in Table I, the relative efficiencies were ranked as PATGL (1st), HAQ-FN (2nd), DOCGL (3rd), PAIN (4th), TJC (5th), CRP (6th), and SJC (7th). Therefore, PATGL, HAQ FN, DOCGL, PAIN, and TJC are the 5 highest ranked measures; PATGL, HAQ-FN, DOCGL, and PAIN are the 4 highest ranked measures; PATGL, HAQ-FN, and DOCGL are the 3 highest ranked measures; and PATGL and HAQ-FN are the 2 highest ranked measures. A simple arithmetic compilation was converted to percentages of instances in which the measure was among the 5, 4, 3, or 2 highest-ranked variables. This approach was selected rather than a direct compilation of the ranks, since differences between relative efficiencies often were quite small. Furthermore, it is not the intention of the authors to rank the measures precisely, as results vary among trials, but rather to gain a sense of whether "objective" measures were clearly superior to other measures to distinguish active from control treatment. The rankings also were compared to rankings of years of disease duration.

A second set of analyses was performed to compare the rankings of only 3 measures, SJC and ESR/CRP (again counted as "one" laboratory test based on the higher rank of ESR or CRP) and HAQ function. The other 4 Core Data Set measures (TJC, DOCGL, Pain, and PATGL) were excluded from these analyses. SJC and ESR/CRP are regarded by most rheumatologists as optimal indicators of inflammation (6). HAQ function is regarded as most likely to be affected by joint damage (21, 22), and therefore the least reversible measure and poorest indicator of inflammation, particularly in patients with long duration of disease. In these analyses, rankings could be only one of the 2 highest variables or highest ranked variable.

Results

Reports studied. Nine comparisons form 8 RA clinical trials of 5 agents, methotrexate, leflunomide, infliximab,

adalimumab, and abatacept, were analysed for their relative efficiencies to distinguish active from control treatments. The median age in all trials was similar, ranging from 50.2 to 57.3 years (Table I). The median duration of disease in all but one trial ranged from 6.5 to 13.1 years (Table I); in one trial, ASPIRE, to compare infliximab to methotrexate, median disease duration was 0.8–0.9 years.

Analyses of relative efficiencies of 7 Core Data Set measures in 9 comparisons

Initial studies of relative efficiencies (7) had been performed in analyses of the US301 clinical trial to compare treatment with leflunomide, methotrexate, and placebo (11). Analyses of these data (Table I) indicated that the highest ranked Core Data Set measure in both the leflunomide *versus* placebo and methotrexate *versus* placebo comparisons was patient global estimate. In the leflunomide *versus* placebo comparison, the second highest rank was seen for HAQ function, compared to physician global estimate in the methotrexate *versus* placebo comparison. The two measures not among the top 5 ranked were CRP/ESR and swollen joint count in the leflunomide *versus* placebo comparison and HAQ-function and swollen joint count in the methotrexate *versus* placebo comparison. No single measure stands out as substantially more or less efficient than other measures, providing a rationale for a 7-measure Core Data Set.

Analyses of infliximab data (Fig. 1) indicated that in ATTRACT (12), the highest relative efficiency was for patient global estimate, followed by pain, SJC, physician global estimate, HAQ-FN, TJC, and CRP (8). In ASPIRE (13) the pattern was quite different, with the highest relative efficiency for SJC, followed by HAQ-FN, physician global estimate, CRP, patient global estimate, pain and TJC (8). Different rankings for the two trials again indicate that no single measure stands out as substantially more or less efficient than other measures.

Analyses of adalimumab data (Table I) with rankings compiled from the published reports of ARMADA (14), DEO111 (16), STAR (17), and DEO19 (15) again indicate variation in the most efficient and least efficient measures but an overall pattern in which no single measure stands out (9). Abatacept data from the ATTAIN trial (Fig. 2) provided the only trial among the 9 comparisons in which highest relative efficiency among RA Core Data Set measures was seen for ESR, although lowest relative efficiency was seen for SJC (10).

Ranking of 7 Core Data Set measures

Data from the 9 comparisons were compiled in Figure 3. The physician global estimate and patient global estimate were among the 5 highest relative efficiencies in all 9 comparisons, followed by HAQ-function and pain in 78%, SJC 67%, ESR/CRP 44% and TJC 33%. Among the 4 highest relative efficiencies, physician global estimate was found in all trials, followed by pain in 78%, HAQ-function and patient global estimate in 67%, ESR/CRP and SJC and 44%, and TJC in none (Fig. 3). Among the 3 highest relative efficiencies, physician global estimate was found in 78% of comparisons, followed by HAQ- function, pain and PATGL in 56%; SJC and ESR/CRP in 33%, and TJC in none. Among the 2 highest relative efficiencies physician and patient global estimates were found in 55% of comparisons, HAQ function in 33%, SJC and pain in 22%, ESR/CRP in 11%, and TJC in none.

HAQ function was ranked 2nd to SJC in the only trial of patients with disease duration of <1 year (ASPIRE). HAQ function also was ranked 2nd in two trials with disease duration of 6.9–7.0 years (US301 methotrexate *vs*. placebo arm), and 10.9–11.0 years (DE019), and 3rd in two trials with mean disease duration of 9.3–11.5 (STAR) and 11.4-12.2 years, (ATTAIN); these 4 comparisons involved the 2nd, 3rd, 4th, and 7th longest median duration of disease among the 9 comparisons (Table I). SJC was ranked 4th, 5th, or 7th in these trials (Table I). ESR was ranked 1st, 6th or 7th in these trials (Table I). Therefore, HAQ appears at least as responsive as SJC or EST to distinguish active from control treatments in patients with median disease duration of 6.9–12.2 years.

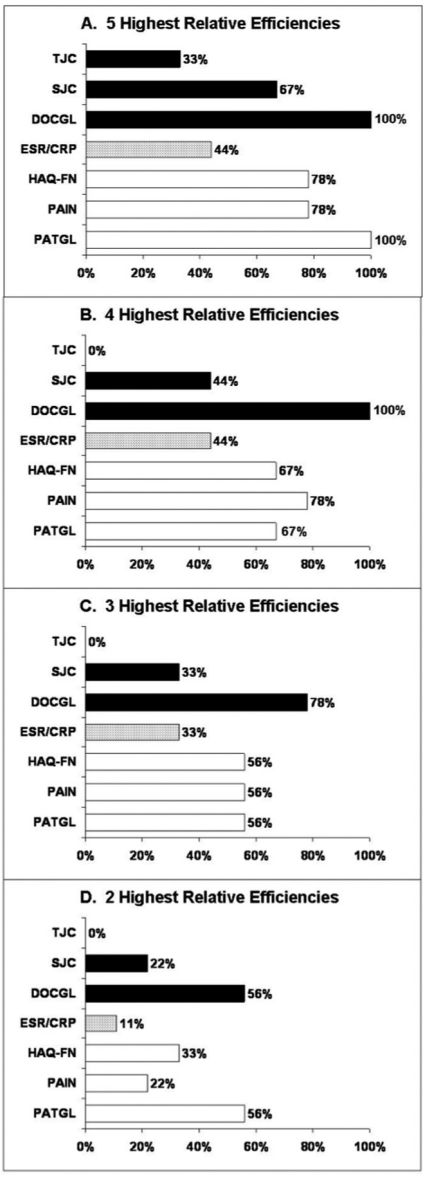

Fig. 3. Ranking of 7 Core Data Set measures in 9 comparisons of active *versus* control treatment. The figure summarises data from Table 1 indicating the number of 9 comparisons in which the 7 Core Data Set measures were among the 5 highest, 4 highest, 3 highest, and 2 highest relative efficiencies. Note that physician global and patient global estimates were among the 5 highest relative efficiencies in all 9 comparisons, HAQ function and pain in 7 of 9 comparisons, SJC in 6 of 7, ESR/CRP in 4 of 7, and TJC in 3 of 7 comparisons. Data indicate the rationale for a core data set of 7 measures as each measure was among the most efficient in 3 comparisons in clinical trials.

Analyses of only 3 Core Data Set measures

A second series of analyses of the 9 trials focused only on SJC, ESR/CRP and HAQ-function (Fig. 4), based on a rationale that SJC and ESR/CRP are regarded by most rheumatologists as optimal indicators of inflammation (6),

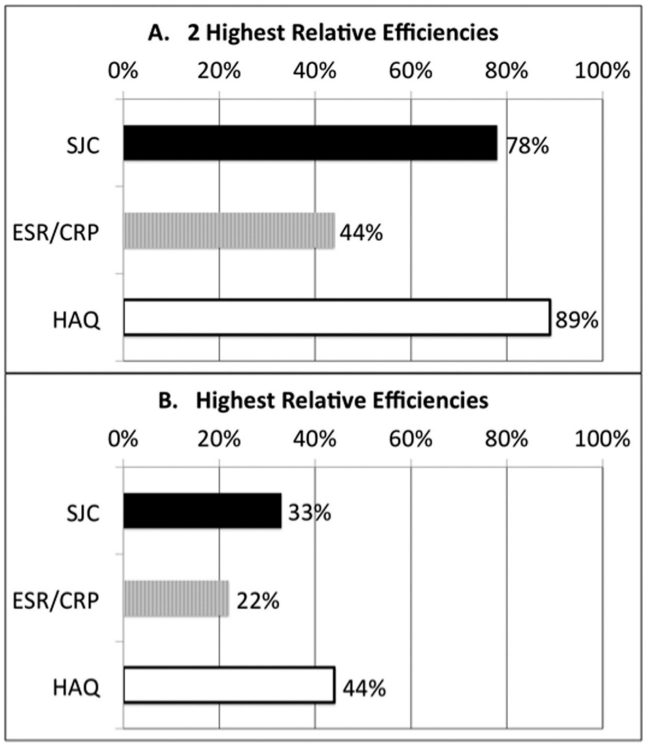

Fig. 4. Relative efficiencies compared only for 3 measures SJC (swollen joint count), ESR/CRP (laboratory data of erythrocyte sedimentation rate/C-reactive protein), and HAQ (Health Assessment Questionnaire) Physical Function score, that indicate that HAQ and SJC are relatively similar in relative efficiencies when considered only among the 3 measures, while laboratory data were efficient in fewer comparisons.

while HAQ function as the least reversible (21, 22) and poorest indicator of inflammation, particularly in patients with long duration of disease. The 2 highest relative efficiencies among the 3 measures were seen for HAQ-function in 8 of the 9 comparisons (89%) (all but ARMADA, and a "tie" with SJC for the US301 methorexate *vs*. placebo comparison); SJC in 7 of the 9 comparisons (78%) (all but US301 leflunomide *vs*. placebo and ATTAIN, and the tie with HAQ function in the US301 methotrexate *vs*. placebo comparison noted above); ESR/CRP in 4 of the 9 (44%). The highest relative efficiency among the 3 measures was seen in 4 of the 9 comparisons for HAQ-function (US301 methotrexate *vs*. placebo arm, DE011, STAR, DE019, in 3 for SJC (ARMADA, ATTRACT, and ASPIRE, and in 2 for the ESR/CRP laboratory test (US301 leflunomide *vs*. placebo and ATTAIN) (Table I, Fig. 4).

Discussion

These data compile and extend available published reports concerning the relative efficiencies of the 7 Core Data Set measures to distinguish active from control treatments in clinical trials. The findings indicate a strong rationale for a Core Data Set of 7 measures, as no single measure was clearly superior in all clinical trials. At the same time, no evidence was seen that "objective" laboratory tests of ESR/CRP, tender joint count, or swollen joint count were superior to "subjective" patient self-report measures or global estimates of the doctor or patient, to differentiate active from control treatments in clinical trials.

Of course, swollen and tender joint counts are far more specific measures for RA than any of the other measures (23), and the presence of swollen joints is required to make a diagnosis of RA. At the same time, swollen and tender joint counts have a number of limitations, primarily in their reliability (24). Indeed, swollen and tender joint counts are more likely to respond to placebo (19) than any of the other Core Data Set measures, reflective of the lower relative efficiency in the US301 clinical trial which included 3 groups of patients who were randomised to methotrexate, leflunomide, or placebo (11).

An elevated ESR or CRP often is a criterion for inclusion of a patient in many clinical trials (25), in large part to be amenable to reduction by therapy and thereby be effective to distinguish active from control treatments. However, most trials specify that only one or the other would be elevated. Almost 50% of RA patients have normal values for ESR or CRP at baseline (26), as mean ESR is considerably lower at this time than in earlier decades (27), though not necessarily recognised in regulatory requirements for access to biological agents. About 20% of patients with elevated ESR have normal CRP and *vice versa* (26), which may result in some patients having little likelihood of meaningful change in the trial. Furthermore, some patients maintain elevated ESR and CRP in the face of clinical improvement, which may also contribute to low relative efficiency (28).

Nonetheless, ESR and/or CRP often constitute the only quantitative clinical data found in the medical records of most patients with RA seen at this time, as standard indices are collected in fewer than 30% of patients in the US (29) (and elsewhere). It appears unfortunate that ESR and CRP are regarded as inadequate for pharmaceutical companies to document the efficacy of active *versus* control treatments in clinical trials, but often are the only quantitative measures available in routine care, despite their limitations in documenting clinical improvement in many patients.

Global estimates of physicians and patients generally were the most efficient of the 7 RA Core Data Set measures to distinguish active from control treatments in the 9 clinical trial comparisons studied. Global estimates may take into account all the information that is not necessarily covered by the information in more specific measures. For example, a physician global estimate may recognise comorbidities, extra-articular disease, and/or damage to joints (even when designed to assess activity), while a patient global estimate may incorporate fatigue or problems with work that are not queried specifically. One limitation of global estimates is non-specificity, *e.g*. a more favourable score may result from a job promotion or financial windfall, or a less favourable score may result from unemployment or severe illness of a relative.

Patient questionnaires also have limitations, including cultural differences,

needs for translation and impact of fibromyalgia, which may cause scores to be elevated in the absence of inflammation (30). At the same time, any index that includes a patient self-report measure, which appears desirable for all rheumatic diseases (not only in RA), will be affected by the presence of psychological distress and fibromyalgia. All measures in clinical care require thoughtful interpretation by a knowledgeable and caring physician, *e.g.* an acute substantial rise of ESR in an RA patient who had a normal ESR three months earlier requires consideration that the cause may be an infection or neoplasm before the elevated ESR can be attributed to a flare of RA.

The data do not address considerations for a measure to be included in an RA index, which does not depend only on statistical criteria. For example, in analyses of different indices composed of various numbers of Core Data Set measures, an index known as RAPID2 which included only patient and physician global estimates distinguished abatacept from control treatment as well as the DAS28, RAPID3 or any other index (31). However, this index has *not* been advocated because of the relative non-specificity of global scores, and the desirability of an index which requires no data from a physician.

Similar considerations of specificity are recognised by the authors that may pertain to inclusion of a joint count and/or laboratory tests in an RA index. The rheumatology community may decide that a formal swollen and/or tender joint count should be given higher weighting than patient self-report or global measures in classification criteria, responses to therapy, remission criteria, etc., on the basis of specificity. However, the absence of a statistical basis for such a policy on the basis of relative specificities in clinical trials must be recognised, although several reports suggest that HAQ disability scores are less reversible than other Core Data Set measures in patients with long duration of disease (21, 22, 32), using analytic techniques different from relative efficiencies. Data presented in this report indicate that HAQ function distinguished active from control treatments as effectively as SJC and ESR/CRP in patients with duration of disease of 6.9–12.2 years. One possible alternative explanation is that SJC and ESR are equally unresponsive to HAQ function in later disease, but further research is needed to clarify these matters.

It has been observed that it makes a major difference whether a patient might have 1 or 11 swollen joints or 2 or 12 swollen joints, but it may not be of major importance whether the patient might have 11 or 12 *versus* 1 or 2, which requires at least 90 seconds to ascertain (24, 33). As noted, collection of RAPID3 at all visits in the infrastructure of clinical care in no way excludes collection of formal joint counts and other measures and indices such as DAS28 or CDAI.

Several limitations should be noted in this study. First, as noted in introductory comments, a meta-analysis could provide a more definitive standard statistical method to compare the different Core Data Set measures. Second, only 9 comparisons have been studied. Third, a number of additional reports concerning anakinra (34), certolizumab (35, 36), secukinumab (37) indicate findings consistent with those described here; however only studies with comparable analyses of relative efficiencies are included. Fourth, data which are similar to findings in this study (as expected) have not been presented to compare indices derived from the individual measures, DAS28, CDAI, and RAPID3, in analyses of trials of adalimumab (38), abatacept (31, 39), certolizumab (35, 36), and infliximab (8), as well as for analyses of remission based on RAPID3 in the ESPOIR cohort (40); the focus of this study was relative efficiencies of the individual measures. Finally, the analyses presented are for groups of patients, as are most analyses of clinical trial data, and individual patients may not follow the patterns of the groups, particularly for which measure might have the highest relative efficiency.

RAPID3 is not a substitute for DAS28, SDAI, or CDAI, which may be collected by a physician or metrologist who regards collection as necessary for optimal care, with no extra work. A careful joint examination, though not necessarily a formal joint count, has always been advocated, together with a patient questionnaire (41). Collection of RAPID3 assures that some quantitative data are available at each visit, with no effort on the part of the physician or metrologist. When a patient completes a questionnaire with RAPID3 in the waiting area, the physician has available quantitative data *before* conversation with the patient, saving time for both doctor and patient, similar to a doctor being aware of blood pressure, status of a healing fracture, etc. before engaging with a patient. It has been observed that it may be "better to have 80% of the information in 100% of patients [than] 100% of the information in 5% of patients" (42).

In conclusion, the 7 Core Data Set measures are of relatively equal efficiency to distinguish active from control treatment in 9 comparisons in 8 clinical trials. There is no statistical advantage to joint counts or laboratory data *versus* patient questionnaire or global estimate data on the basis of relative efficiencies. The most specific measures are not necessarily the most sensitive or efficient measures to distinguish active from placebo treatments. Furthermore, patient self-report questionnaire data are the most feasible and cost-effective measures in terms of time and resources for the medical system (43). This information might be considered in recommendations, particularly for routine care in busy clinical settings.

Acknowledgements

We thank Drs Joel Block, Kathryn A. Gibson, Gary G. Koch, Josef Smolen, and Yusuf Yazici, for thoughtful comments concerning this manuscript.

References

1. FELSON DT, ANDERSON JJ, BOERS M *et al*.: The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid Arthritis Clinical Trials. *Arthritis Rheum* 1993; 36: 729-40.
2. CASTREJON I, MCCOLLUM L, TANRIOVER MD, 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* (Hoboken) 2012; 64: 1250-5.
3. 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: 687-97, v.
4. ENGEL GL: The need for a new medical model: a challenge for biomedicine. *Science* 1977; 196: 129-36.
5. FELSON DT, ANDERSON JJ, BOERS M *et al*.: American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995; 38: 727-35.
6. FELSON DT, SMOLEN JS, WELLS G *et al*.: American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Arthritis Rheum* 2011; 63: 573-86.
7. TUGWELL P, WELLS G, STRAND V *et al*.: Clinical improvement as reflected in measures of function and health-related quality of life following treatment with leflunomide compared with methotrexate in patients with rheumatoid arthritis: sensitivity and relative efficiency to detect a treatment effect in a twelve-month, placebo-controlled trial. Leflunomide Rheumatoid Arthritis Investigators Group. *Arthritis Rheum* 2000; 43: 506-14.
8. PINCUS T, ZELINGER D, BOLCE RJ, BAKER D: Relative efficiencies of 7 RA Core Data Set measures and 3 indices, DAS28 (Disease Activity Score), CDAI (Clinical Disease Activity Index) and RAPID3 (Routine Assessment of Patient Index Data), to distinguish infliximab from control treatments in the Attract and Aspire clinical Trials. *Ann Rheum Dis* 2009; 68: 551.
9. PINCUS T, AMARA I, SEGURADO OG, BERGMAN M, KOCH GG: Relative efficiencies of physician/assessor global estimates and patient questionnaire measures are similar to or greater than joint counts to distinguish adalimumab from control treatments in rheumatoid arthritis clinical trials. *J Rheumatol* 2008; 35: 201-5.
10. WELLS G, LI T, MAXWELL L, MACLEAN R, TUGWELL P: Responsiveness of patient reported outcomes including fatigue, sleep quality, activity limitation, and quality of life following treatment with abatacept for rheumatoid arthritis. *Ann Rheum Dis* 2008; 67: 260-5.
11. STRAND V, COHEN S, SCHIFF M *et al*.: Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate. Leflunomide Rheumatoid Arthritis Investigators Group. *Arch Intern Med* 1999; 159: 2542-50.
12. MAINI R, ST CLAIR EW, BREEDVELD F *et al*.: Infliximab (chimeric anti-tumour necrosis factor alpha monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomised phase III trial. ATTRACT Study Group. *Lancet* 1999; 354: 1932-9.
13. ST CLAIR EW, VAN DER HEIJDE DM, SMOLEN JS *et al*.: Combination of infliximab and methotrexate therapy for early rheumatoid arthritis: a randomized, controlled trial. *Arthritis Rheum* 2004; 50: 3432-43.
14. WEINBLATT ME, KEYSTONE EC, FURST DE *et al*.: Adalimumab, a fully human anti-tumor necrosis factor alpha monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate: the ARMADA trial. *Arthritis Rheum* 2003; 48: 35-45.
15. KEYSTONE EC, KAVANAUGH AF, SHARP JT *et al*.: Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled, 52-week trial. *Arthritis Rheum* 2004; 50: 1400-11.
16. VAN DE PUTTE LB, RAU R, BREEDVELD FC *et al*.: Efficacy and safety of the fully human anti-tumour necrosis factor alpha monoclonal antibody adalimumab (D2E7) in DMARD refractory patients with rheumatoid arthritis: a 12 week, phase II study. *Ann Rheum Dis* 2003; 62: 1168-77.
17. FURST DE, SCHIFF MH, FLEISCHMANN RM *et al*.: Adalimumab, a fully human anti tumor necrosis factor-alpha monoclonal antibody, and concomitant standard antirheumatic therapy for the treatment of rheumatoid arthritis: results of STAR (Safety Trial of Adalimumab in Rheumatoid Arthritis). *J Rheumatol* 2003; 30: 2563-71.
18. GENOVESE MC, BECKER JC, SCHIFF M *et al*.: Abatacept for rheumatoid arthritis refractory to tumor necrosis factor alpha inhibition. *N Engl J Med* 2005; 353: 1114-23.
19. STRAND V, COHEN S, CRAWFORD B, SMOLEN JS, SCOTT DL, LEFLUNOMIDE INVESTIGATORS GROUPS: Patient-reported outcomes better discriminate active treatment from placebo in randomized controlled trials in rheumatoid arthritis. *Rheumatology* (Oxford) 2004; 43: 640-7.
20. WARE JE, JR., SHERBOURNE CD: The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical care* 1992; 30: 473-83.
21. WELSING PM, VAN GESTEL AM, SWINKELS HL, KIEMENEY LA, VAN RIEL PL: The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. *Arthritis Rheum* 2001; 44: 2009-17.
22. ALETAHA D, SMOLEN J, WARD MM: Measuring function in rheumatoid arthritis: Identifying reversible and irreversible components. *Arthritis Rheum* 2006; 54: 2784-92.
23. PINCUS T: The DAS is the most specific measure, but a patient questionnaire is the most informative measure to assess rheumatoid arthritis. *J Rheumatol* 2006; 33: 834-7.
24. SOKKA T, PINCUS T: Joint counts to assess rheumatoid arthritis for clinical research and usual clinical care: advantages and limitations. *Rheum Dis Clin North Am* 2009; 35: 713-22, v-vi.
25. SOKKA T, PINCUS T: Most patients receiving routine care for rheumatoid arthritis in 2001 did not meet inclusion criteria for most recent clinical trials or american college of rheumatology criteria for remission. *J Rheumatol* 2003; 30: 1138-46.
26. 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: 1387-90.
27. ABELSON B, SOKKA T, PINCUS T: Declines in erythrocyte sedimentation rates in patients with rheumatoid arthritis over the second half of the 20th century. *J Rheumatol* 2009; 36: 1596-9.
28. WOLFE F, PINCUS T: The level of inflammation in rheumatoid arthritis is determined early and remains stable over the longterm course of the illness. *J Rheumatol* 2001; 28: 1817-24.
29. ANDERSON J, CAPLAN L, YAZDANY J *et al*.: Rheumatoid arthritis disease activity measures: American College of Rheumatology recommendations for use in clinical practice. *Arthritis Care Res* (Hoboken) 2012; 64: 640-7.
30. PINCUS T, YAZICI Y, BERGMAN MJ: Patient questionnaires in rheumatoid arthritis: advantages and limitations as a quantitative, standardized scientific medical history. *Rheum Dis Clin North Am* 2009; 35: 735-43, vii.
31. 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: 345-9.
32. ALETAHA D, ALASTI F, SMOLEN JS: Chronicity of rheumatoid arthritis affects the responsiveness of physical function, but not of disease activity measures in rheumatoid arthritis clinical trials. *Ann Rheum Dis* 2014.
33. PINCUS T, SWEARINGEN CJ, BERGMAN MJ *et al*.: RAPID3 (Routine Assessment of Patient Index Data) on an MDHAQ (Multidimensional Health Assessment Questionnaire): agreement with DAS28 (Disease Activity Score) and CDAI (Clinical Disease Activity Index) activity categories, scored in five versus more than ninety seconds. *Arthritis Care Res* (Hoboken) 2010; 62: 181-9.
34. COHEN SB, STRAND V, AGUILAR D, OFMAN JJ: Patient- versus physician-reported outcomes in rheumatoid arthritis patients treated with recombinant interleukin-1 receptor antagonist (anakinra) therapy. *Rheumatology* (Oxford) 2004; 43: 704-11.
35. PINCUS T, FURER V, KEYSTONE E, YAZICI Y, BERGMAN MJ, LUIJTENS K: RAPID3 (Routine Assessment of Patient Index Data 3) severity categories and response criteria: Similar results to DAS28 (Disease Activity Score) and CDAI (Clinical Disease Activity Index) in the RAPID 1 (Rheumatoid Arthritis Prevention of Structural Damage) clinical trial of certolizumab pegol. *Arthritis Care Res* (Hoboken) 2011; 63: 1142-9.
36. STRAND V, SMOLEN JS, VAN VOLLENHOVEN RF *et al*.: Certolizumab pegol plus methotrexate provides broad relief from the burden of rheumatoid arthritis: analysis of patient-reported outcomes from the RAPID 2 trial. *Ann Rheum Dis* 2011; 70: 996-1002.
37. STRAND V, KOSINSKI M, GNANASAKTHY A, MALLYA U, MPOFU S: Secukinumab treatment in rheumatoid arthritis is associated with incremental benefit in the clinical outcomes and HRQoL improvements that ex-

ceed minimally important thresholds. *Health and quality of life outcomes* 2014; 12: 31.

38. PINCUS T, CHUNG C, SEGURADO OG, AMARA I, KOCH GG: An index of patient reported outcomes (PRO-Index) discriminates effectively between active and control treatment in 4 clinical trials of adalimumab in rheumatoid arthritis. *J Rheumatol* 2006; 33: 2146-52.
39. PINCUS T, HINES P, BERGMAN MJ, YAZICI Y, ROSENBLATT LC, MACLEAN R: Proposed severity and response criteria for Routine Assessment of Patient Index Data (RAPID3): results for categories of disease activity and response criteria in abatacept clinical trials. *J Rheumatol* 2011; 38: 2565-71.
40. CASTREJON I, DOUGADOS M, COMBE B, GUILLEMIN F, FAUTREL B, PINCUS T: Can remission in rheumatoid arthritis be assessed without laboratory tests or a formal joint count? possible remission criteria based on a self-report RAPID3 score and careful joint examination in the ESPOIR cohort. *J Rheumatol* 2013; 40: 386-93.
41. PINCUS T, GIBSON KA, BERTHELOT JM:, Is a patient questionnaire without a joint examination as undesirable as a joint examination without a patient questionnaire? *J Rheumatol* 2014; 41: 619-21.
42. PINCUS T, WOLFE F: Patient questionnaires for clinical research and improved standard patient care: is it better to have 80% of the information in 100% of patients or 100% of the information in 5% of patients? *J Rheumatol* 2005; 32: 575-7.
43. PINCUS T, SKUMMER PT, GRISANTI MT, CASTREJON 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: 177-86.