53

Original article

GUEPARD treat-to-target strategy is significantly more efficacious than ESPOIR routine care in early rheumatoid arthritis according to patient-reported outcomes and physician global estimate

Isabel Castrejón[1], Theodore Pincus[1], Martin Soubrier[2], Yih Chang Lin[3], Anne-Christine Rat[4], Bernard Combe[5] and Maxime Dougados[6]

Abstract

Objective. To analyse seven RA Core Data Set measures and three indices for their capacity to distinguish treatment results in early RA in the GUEPARD treat-to-target clinical trial *vs* ESPOIR routine care.

Methods. *Post hoc* analyses compared 65 GUEPARD and 130 matched control ESPOIR patients over 6 and 12 months for mean changes in measures, relative efficiencies and standardized response means (SRM). Three indices—28-joint disease activity score (DAS28), clinical disease activity index (CDAI) and routine assessment of patient index data (RAPID3)—were compared for mean changes and numbers of patients with high, moderate or low activity or remission using κ values.

Results. Greater improvement was seen for GUEPARD *vs* ESPOIR, statistically significant for physician and patient global estimates and pain and health assessment questionnaire physical function (HAQ-FN), but not joint counts and laboratory tests. Relative efficiencies with tender joint count as the referent measure indicated that pain (2.57) and global estimates by patient (3.13) and physician (2.31) were most efficient in distinguishing GUEPARD from ESPOIR. Mean improvements in GUEPARD *vs* ESPOIR were -3.4 *vs* -2.6 for DAS28 (0–10) (24%), -29.8 *vs* -23.1 for CDAI (0–76) (23%) and -13.0 *vs* -7.8 for RAPID3 (0–30) (40%) (all $P < 0.01$); agreement was moderate between CDAI *vs* DAS28 ($\kappa = 0.56$) and *vs* RAPID3 ($\kappa = 0.48$), and fair between DAS28 *vs* RAPID3 ($\kappa = 0.26$).

Conclusion. Patient and global measures indicate greater efficacy than joint counts or laboratory measures in detecting difference between GUEPARD treat-to-target and ESPOIR routine care. A RAPID3 of only patient measures may help guide treat-to-target in busy clinical settings.

Key words: treat-to-target, patient-reported outcomes, assessment, rheumatoid arthritis, patient questionnaires.

[1]Department of Medicine, Division of Rheumatology, New York University School of Medicine and NYU Hospital for Joint Diseases, New York, NY, USA, [2]Department of Rheumatology, G. Montpied Hospital, Clermont-Ferrand, France, [3]Department of Rheumatic and Immunologic Disease, Cleveland Clinic, Cleveland, OH, USA, [4]Université de Lorraine, Paris Descartes University, APEMAC, EA 4360, Nancy, [5]University Hospital, Department of Rheumatology, Montpellier and [6]Department of Rheumatology, Cochin Hospital, Paris, France.

Introduction

RA is characterized by the absence of a single gold standard measure, such as blood pressure or serum haemoglobin A1c, that can be applied to all individual patients with a specific diagnosis [1]. Therefore a core data set of eight measures and indices of multiple measures have been developed for patient assessment [2]. The RA Core Data Set includes three measures from a health professional [tender joint count (TJC), swollen joint count (SJC), physician global estimate of status (DOCGL)]; three from a patient self-report questionnaire [physical function (FN), pain

and patient global estimate of status (PATGL)]; one laboratory test of an acute phase reactant, either ESR or CRP; and, in studies of 12 months or longer, a radiographic score.

Indices of multiple measures from the Core Data Set have been developed to assess and monitor RA patients in clinical trials and clinical care. The 28-joint DAS (DAS28) is the most specific index for RA [3], but it has limitations in routine care, including complex calculations (albeit available at an excellent website, http://www.4s-dawn.com/DAS28/DAS28.html); the need for a laboratory test, which often is not available at the time of the visit and in some situations is poorly informative [4], and the need for a formal joint count. The clinical disease activity index (CDAI) [5, 6] simplifies calculations and does not require a laboratory test, but does require formal joint counts, which have limitations in measurement [7] and usually are not performed at most visits of RA patients [8]. Routine assessment of patient index data (RAPID3), an index of only the three patient self-report Core Data Set measures, distinguishes active from control treatments in clinical trials of LEF [9], MTX [9], adalimumab [10], abatacept [11] and certolizumab [12] at levels similar to DAS28 and CDAI, and requires only about 5 s to score on a multidimensional health assessment questionnaire (MDHAQ) *vs* more than 100 s to score a DAS28 or CDAI [13].

A treat-to-target strategy guided by DAS28 results in better outcomes than traditional non-quantitative care [14]. A recent report compared results of treatment in two populations of early RA patients: the GUEPARD (Guérir la Polyarthrite Rhumatoïde Débutante) clinical trial with treat-to-target tight control based on DAS28 *vs* the ESPOIR (Etude et Suivi des Polyarthrites Indifférenciées Récentes) cohort with routine care [15]. Improvement appeared greater in GUEPARD *vs* ESPOIR patients, at higher levels according to patient and global RA Core Data Set measures than according to joint counts and laboratory tests. These observations suggested that further formal comparisons of each RA Core Data Set measure and DAS28, CDAI and RAPID3 to distinguish results of GUEPARD tight control *vs* ESPOIR routine care would be of value, as presented in this report.

Materials and methods

Databases

Post hoc analyses were performed on two databases of patients with early RA: 65 patients from the GUEPARD tight control trial and 130 matched patients from the ESPOIR usual care cohort [15]. GUEPARD includes patients with RA as defined by the 1987 criteria of the ACR with maximum disease duration of 6 months and DAS28-ESR >5.1. Patients who met eligibility criteria were randomized to receive MTX monotherapy or a combination of MTX and adalimumab (ADA), with adjustment of the treatment every 3 months to achieve a DAS28-ESR <3.2. If this target was not met, the following steps were taken

sequentially: MTX and ADA (40 mg every other week), MTX and ADA (40 mg/week), MTX and etanercept (25 mg twice a week) and MTX combined with LEF.

ESPOIR is a large national multicentre, longitudinal and prospective cohort initiated by the French Society of Rheumatology [16]. Patients with undifferentiated arthritis or RA, with disease duration <6 months who were both DMARD- and glucocorticoid-naïve, were recruited. Patients were seen every 6 months during the 2 first years and then every year. The patients were treated with routine care, without a specific target, and followed by their rheumatologists, without a treatment protocol for a target. The approach was determined entirely by the treating rheumatologist, including a decision to initiate biologic agents. These prior studies and compilation of the databases for further analyses were conducted with the approval of the ESPOIR central ethics committee of Montpellier and the GUEPARD central ethics committee of Cochin, Paris. The analysis plan for this study was submitted to the Institutional Review Board (IRB) of the New York University School of Medicine, which deemed the planned analyses of deidentified data exempt from requirement of IRB approval.

A total of 130 patients in the ESPOIR cohort fulfilling the inclusion criteria of the GUEPARD trial and with no missing data at baseline were matched 1:2 using a propensity score. A propensity score is used in statistical analyses of non-randomized data to attempt to match patients to provide an unbiased estimation of group effects by adjusting for specific variables. The main advantage of a propensity score over random assignment is that it avoids the ethical considerations that arise when a potentially beneficial treatment or strategy is denied at random. Non-randomized data are less costly than clinical trials, and a propensity score allows the analysis of existing data. A propensity score is limited, as one controls only for recognized variables, while hidden bias may remain because of other non-recognized variables. Random assignment provides more confidence that both groups are similar on both observed and non-observed characteristics. In addition, matching can only estimate effects where there is overlap between groups, whereas random assignment ensures that there is common support across the whole sample.

The propensity score was computed using a multivariate logistic model. The following demographic and disease characteristics at baseline were used as covariates in the model: treatment centre; gender; age at inclusion; disease duration; BMI; RF and anti-CCP antibody status; ESR; CRP; tender and swollen 28 joint counts; patient's assessments of pain, disease activity and fatigue; doctor's assessment of disease activity on a 0–100 mm visual analogue scale (VAS); health assessment questionnaire physical function (HAQ-FN); physical and mental components of the 36-item Short Form Health Survey (SF-36); erosive disease, erosion score and narrowing score. A more detailed description of the matching methodology and these patient groups is found in a previous report [15].

Study RA Core Data Set measures

The RA Core Data Set measures include the number of tender joints (TJC) (0–28); number of swollen joints (SJC) (0–28); physician global estimate (DOCGL) (VAS 0–100); HAQ-FN (0–3); pain (VAS 0–100); patient global estimate (PATGL) (VAS 0–100); ESR (mm/h) or CRP (mg/dl); and, in studies of ⩾ 12 months, radiographic scores according to the modified Sharp/van der Heijde score. Each of these variables was analysed at baseline, and at 6 and 12 months of follow-up. HAQ-FN was recalculated from 0–3 to 0–10 (multiplied by 3.3) in order to score RAPID3 (see below).

Composite RA indices

Three composite indices—DAS28, CDAI and RAPID3—were analysed at baseline, 6 months and 12 months. DAS28 is calculated according to the formula: $DAS28 = (0.56 \times \sqrt{TJC}) + (0.28 \times \sqrt{SJC}) + [0.7 \times \ln (ESR)] + [0.014 \times PATGL \text{ (in mm)}]$, for a total score of 0–10 [3]. DAS28 categories include remission (⩽ 2.6), low activity (2.6–3.2), moderate activity (3.2–5.1) and high activity (>5.1) (Supplementary data Table S1, available at *Rheumatology* Online).

CDAI is a simple sum of TJC (0–28) + SJC (0–28) + DOCGL (0–10) + PATGL (0–10), for a total score range of 0–76 [6]. CDAI categories include remission (⩽ 2.8), low activity (2.9–10.0), moderate activity (10.1–22.0) and high activity (>22) (Supplementary data Table S1, available at *Rheumatology* Online).

RAPID3 is a composite index of three 0–10 scales for HAQ-FN (recalculated from 0–3 on the HAQ, as noted above), pain and PATGL, for a total score range of 0–30 [12]. RAPID3 categories include remission (⩽ 3), low activity (3.1–6.0), moderate activity (6.1–12.0) and high activity (>12) [13] (Supplementary data Table S1, available at *Rheumatology* Online).

Statistical analysis

All statistical analyses were performed using Stata 12.0 for Windows (StataCorp LP, College Station, TX, USA). Baseline demographic and clinical measures were compared in GUEPARD tight control *vs* ESPOIR routine care using the Student's *t* test for variables normally distributed and Wilcoxon rank-sum test for variables with non-normal distribution.

Differences between GUEPARD (tight control) *vs* ESPOIR (routine care) in the responsiveness of each measure to document clinical changes between baseline and 12 months were calculated according to the method of Tugwell *et al.* [17]. Initially, standardized response means (SRMs) were calculated for each measure, as described by Wells *et al.* [18] (see supplementary data Appendix for formulas, available at Rheumatology Online). Relative efficiencies are expressed in relation to the TJC as the referent measure [17].

A relative efficiency >1 indicates that the measure is more efficient than TJC to detect a difference between results of two treatment strategies. These analyses were performed without adjustment for mean daily dosage of corticosteroids or the duration of biological agent therapy, as in the previous report [15]. Results were similar at 6 and 12 months, and only 12-month data are presented for analyses of individual measures.

Each of the three composite indices—DAS28, CDAI and RAPID3—was calculated at baseline, 6 months and 12 months. Mean changes from baseline to endpoints at 6 months and 12 months were compared in GUEPARD *vs* ESPOIR using Student's *t* test, with statistical significance adjusted for eight comparisons for individual measures and three comparisons for indices. The numbers and proportions of patients in four categories—high, moderate and low activity and remission—were computed at baseline, 6 and 12 months according to each index. The agreement of the disease activity categories was compared using κ statistics [19]; strength of agreement was interpreted as κ >0.81 = very good agreement, 0.61–0.80 = substantial agreement, 0.41–0.60 = moderate agreement, 0.21–0.40 = fair agreement and ⩽ 0.20 = poor agreement [20].

Results

Individual RA Core Data Set measures

Baseline demographic and disease variables were similar in GUEPARD and ESPOIR (Table 1). The only significant differences were that patients in the GUEPARD tight control clinical trial had longer mean disease duration *vs* the ESPOIR routine care patients (5.4 *vs* 3.5 months), higher likelihood of treatment with biologic therapies (82% *vs* 13%), a lower percentage taking glucocorticoids (50% *vs* 75%) and a lower mean dose of glucocorticoids (4.7 *vs* 8.1) (all $P < 0.001$) (Table 1).

Substantial improvement was seen in both GUEPARD and ESPOIR after 6 and 12 months of follow-up (Table 2). Differences in improvement were statistically significantly greater in the GUEPARD tight control group for HAQ-FN, pain, PATGL and TJC unadjusted, and all but TJC adjusted for eight comparisons. However, differences were not statistically significantly greater in GUEPARD *vs* ESPOIR for SJC, ESR, CRP and radiographic progression (Table 2).

High SRMs >4 were observed for PATGL, pain and DOCGL, compared with 2–3 for TJC and HAQ-FN, and <1 for SJC, ESR, CRP and radiographic score (Table 3). Relative efficiencies with TJC as the referent measure indicated that PATGL (3.13), pain (2.57) and DOCGL (2.31) were more efficient than TJC to distinguish GUEPARD from ESPOIR (Table 3), while HAQ-FN (0.43), SJC (0.41), ESR (0.01), CRP (0.12) and radiographic score (0.31) were less efficient than TJC to recognize differences between the two patient groups (Table 3).

Three RA indices: DAS28, CDAI and RAPID

Statistically significant improvements from mean baseline values over 6 and 12 months were seen in both GUEPARD and ESPOIR, according to DAS28, CDAI and RAPID3 (Table 4). Mean improvement over 6 and 12 months was

TABLE 1 Baseline demographic and clinical characteristics of patients in GUEPARD (tight control) and ESPOIR (routine care)

	GUEPARD (n = 65) tight control	ESPOIR (n = 130) routine care	P-value
Demographic variables			
Age, mean (S.D.), years	47.9 (15.7)	48.1 (11.5)	0.91
Women, n (%)	52 (80.0)	98 (75.4)	0.59
Core Data Set measures			
TJC (0–28), mean (S.D.)	14.0 (6.7)	14.1 (6.8)	0.95[a]
SJC (0–28), mean (S.D.)	10.1 (4.9)	10.7 (5.7)	0.75[a]
Physician global assessment (VAS 0–100), mean (S.D.)	67.6 (17.2)	65.9 (18.1)	0.53
HAQ-FN score adjusted (0–3 scale), mean (S.D.)	1.5 (0.6)	1.5 (0.7)	0.23
Pain score (VAS 0–100), mean (S.D.)	59.5 (22.0)	55.7 (20.7)	0.23
Patient global assessment (VAS 0–100), mean (S,D,)	68.1 (18.7)	67.9 (21.9)	0.86[a]
ESR, mean (S.D.) mm/h	37.3 (22.4)	39.5 (25.7)	0.76[a]
CRP, mean (S.D.), mg/l	28.7 (33.4)	29.8 (32.3)	0.36[a]
Modified Sharp–van der Heijde (0–448), mean (S.D.)	7.7 (13.4)	5.8 (8.0)	0.70[a]
Erosion score (0–280), mean (S.D.)	2.4 (4.9)	2.4 (3.7)	0.10[a]
JSN score (0–168), mean (S.D.)	5.3 (9.5)	3.5 (5.4)	0.11[a]
RA indices			
DAS28-ESR (0–10), mean (S.D.)	6.2 (0.8)	6.3 (0.9)	0.77[a],*
DAS28-CRP (0–10), mean (S.D.)	5.7 (1.2)	5.9 (1.1)	0.74
CDAI (0–76), mean (S.D.)	37.8 (10.9)	38.2 (12.8)	0.81
RAPID3 (0–30), mean (S.D.)	17.9 (5.2)	17.3 (5.1)	0.45
Other disease measure			
Symptom duration, mean (S.D.), months	5.4 (4.7)	3.5 (2.0)	<0.001[a]
Therapy variables			
Patients receiving biologic therapy during the study, n (%)	53 (82)	17 (13)	<0.0001*
Duration of biologic therapy, mean (S.D.), months	6.8 (3.6)	4.9 (3.6)	0.10
Patients receiving glucocorticoid treatment, n (%)	33 (50)	97 (75)	0.001*
Daily dose of prednisolone, mean (S.D.), mg	4.7 (3.3)	8.1 (5.2)	0.0008*

[a]Wilcoxon rank-sum test for non-normally distributed variables. *Statistically significant results at $P < 0.05$. Student's *t*-test for normally distributed variables.

TABLE 2 Changes in variables over 12 months in GUEPARD (tight control) *vs* ESPOIR (routine care)

	GUEPARD (N = 65) tight control, mean change (95% CI)	ESPOIR (N = 130) routine care, mean change (95% CI)	P-value[a]	Adjusted P-value
TJC (0–28)	−11.2 (−9.4, −13)	−8.7 (−10.1, −7.2)	0.03	0.27
SJC (0–28)	−8.2 (−9.5, −6.9)	−7.7 (−8.8, −6.6)	0.60	NP
Physician global assessment (VAS 0–100)	−51.5 (−57.6, −45.2)	−34.7 (−39.6, −29.9)	<0.001*	<0.009*
HAQ physical function score adjusted (0–3)	−0.97 (−1.31, −0.82)	−0.66 (−0.78, −0.54)	0.003*	0.027*
Pain score (VAS 0–100)	−44.8 (−51.1, −38.6)	−24.2 (−28.7, −19.6)	<0.001*	<0.009*
Patient global assessment (VAS 0–100)	−50.2 (−56.5, −43.9)	−30.5 (−35.9, −25.1)	<0.001*	<0.009*
ESR, mm/h	−20.4 (−25.6, −15.2)	−20.2 (−24.6, −15.8)	0.76	NP
CRP, mg/l	−22.5 (−30.8, −14.1)	−19.9 (−25.9, −13.9)	0.97	NP
Modified Sharp–van der Heijde score (0–448 scale)	1.1 (−0.2, 1.9)	1.4 (0.8, 2.1)	0.25	NP

NP: not performed (adjusted *P*-value not calculated when unadjusted *P*-value >0.05). [a]Student's *t*-test. *Statistically significant results at $P < 0.005$ (Bonferroni adjustment for nine comparisons).

significantly greater in GUEPARD *vs* ESPOIR. Differences at 6 months were −3.3 *vs* −2.2 (33.3%) for DAS28 (0–10), −29.5 *vs* −20.9 (29.2%) for CDAI (0–76) and −12.1 *vs* −7.2 (40.5%) for RAPID3 (0–30) (all *P* <0.001).

Differences at 12 months were −3.4 *vs* −2.6 (23.5%) for DAS28 (0–10), −29.8 *vs* −23.1 (22.4%) for CDAI (0–76) and −13.0 *vs* −7.8 (40.0%) for RAPID3 (0–30) (all *P* <0.01) (Table 4).

TABLE 3 Differences in individual measures to distinguish change over 12 months in GUEPARD (tight control strategy) *vs* ESPOIR (routine care)

Variable	Standardized response mean (95% CI)	Relative efficiency[a]
TJC (0–28)	2.15 (0.18, 4.12)	—[a]
SJC (0–28)	0.52 (−1.45, 2.49)	0.24
Physician global assessment (VAS 0–100)	4.09 (2.12, 6.10)	1.90
HAQ physical function score adjusted (0–3)	2.94 (0.97, 4.92)	1.37
Pain score (VAS 0–100)	5.24 (3.27, 7.22)	2.43
Patient global assessment (VAS 0–100)	4.43 (2.46, 6.41)	2.06
ESR, mm/h	0.06 (−2.03, 1.91)	0.03
CRP, mg/l	0.48 (−2.45, 1.48)	0.22
Modified Sharp–van der Heijde score (0–448 scale)	0.67 (−2.64, 1.30)	0.31
Erosion score (0–280 scale)	1.72 (0.25, 3.69)	0.80
JSN score (0–168 scale)	2.07 (0.10, 4.04)	0.96

[a]Relative efficiency of each measure (compared with TJC as the referent measure) to distinguish GUEPARD *vs* ESPOIR according to the method of Tugwell *et al.* [17]. JSN: joint space narrowing.

TABLE 4 Mean changes in three indices over 6 and 12 months in GUEPARD[a] and ESPOIR[b]

	GUEPARD (*N* = 65) tight control		ESPOIR (*N* = 130) routine care		
	Mean (s.d.)	Mean change from baseline (s.d.)	Mean (s.d.)	Mean change from baseline (s.d.)	P value[c]
DAS28 (0–10)					
Baseline	6.25 (0.10)	NA	6.27 (0.08)	NA	—
6 months	2.97 (0.15)	−3.29 (1.33)	4.00 (0.13)	−2.25 (1.50)	<0.0001
12 months	2.92 (0.16)	−3.36 (1.25)	3.72 (0.14)	−2.57 (1.61)	0.0008
CDAI (0–76)					
Baseline	37.80 (1.4)	NA	38.23 (1.13)	NA	—
6 months	8.37 (1.08)	−29.47 (12.88)	17.06 (1.11)	−20.86 (15.01)	0.0001
12 months	8.22 (1.23)	−29.76 (13.01)	15.28 (1.21)	−23.06 (16.37)	0.005
RAPID3 (0–30)					
Baseline	17.90 (0.65)	NA	17.34 (0.45)	NA	—
6 months	5.40 (0.69)	−12.09 (5.89)	10.07 (0.56)	−7.24 (6.89)	<0.0001
12 months	4.78 (0.72)	−12.95 (5.90)	9.93 (0.63)	−7.77 (6.61)	<0.0001

[a]GUEPARD tight control strategy. [b]ESPOIR routine care. [c]Student's *t*-test comparing mean change from baseline to 6 or 12 months between the two cohorts.

The percentage of patients in remission after 6 months was significantly higher in the GUEPARD tight control group and similar for the three indices (Table 5). At 6 months, moderate agreement was seen between CDAI and DAS28 ($\kappa = 0.56$) and between RAPID3 and CDAI ($\kappa = 0.48$); fair agreement was seen between DAS28 and RAPID3 ($\kappa = 0.26$). At 12 months, substantial agreement was seen between CDAI and DAS28 ($\kappa = 0.62$); fair agreement was seen between RAPID3 and CDAI ($\kappa = 0.37$) and between DAS28 and RAPID3 ($\kappa = 0.31$).

Discussion

This report presents data consistent with extensive evidence that a treat-to-target strategy has greater efficacy in the treatment of RA than routine care [14]. To the best knowledge of the authors, this is the first report to compare relative efficiencies of RA Core Data Set measures in two independently recruited patient groups rather than two arms of a clinical trial. The greater efficacy of the GUEPARD tight control treat-to-target strategy *vs* ESPOIR routine care is seen at higher levels according to patient self-report and global measures compared with joint counts and laboratory tests.

These findings also are consistent with evidence of high relative efficiencies of patient report measures compared with joint counts and laboratory tests to distinguish active from control treatments in clinical trials of LEF [9], MTX [9], adalimumab [10], abatacept [11] and certolizumab [12]. In these trials, variation in the most efficient measure is seen, although DOCGL, PATGL, pain and HAQ-FN generally

TABLE 5 Percentage of patients in four activity categories at 6 months and 12 months in GUEPARD (tight control) and ESPOIR (routine care)

Cohorts	Composite index	% of patients in four activity categories at 6 months			
		High	Moderate	Low	Remission
GUEPARD (*n* = 65) tight control	DAS28 (0–10)	8	34	9	49
	CDAI (0–76)	11	23	28	38
	RAPID3 (0–30)	16	16	30	38
ESPOIR (*n* = 130) routine care	DAS28 (0–10)	23	45	16	16
	CDAI (0–76)	30	37	20	13
	RAPID3 (0–30)	40	28	14	18
Cohorts	Composite index	% of patients in four activity categories at 12 months			
		High	Moderate	Low	Remission
GUEPARD (*n* = 65) tight control	DAS28 (0–10)	6	31	15	48
	CDAI (0–76)	9	22	28	41
	RAPID3 (0–30)	11	17	19	53
ESPOIR (*n* = 130) routine care	DAS28 (0–10)	20	36	16	28
	CDAI (0–76)	27	29	29	15
	RAPID3 (0–30)	41	26	10	23

show greater efficiency than TJC, SJC, ESR and CRP. However, each measure is most efficient in a certain trial, underscoring a need for an RA Core Data Set of seven measures. Each RA Core Data Set measure has a similar order of magnitude to detect differences in clinical trials and clinical care. Patient self-report and global measures have greater sensitivity, while joint counts and laboratory tests have greater specificity.

RAPID3 is an index that includes only self-report data provided by patients, without a need for formal joint counts from health professionals, scored in 5 s on a multidimensional health assessment questionnaire (MDHAQ) *vs* ⩾100 s for DAS28 or CDAI [13]. RAPID3 yielded similar, but not identical, results compared to DAS28 or CDAI for categories of high, moderate and low activity, and remission in this study and others [13, 21]. All RAPID3 data are from the same observer—the patient—overcoming the need for the same health professional observer to perform a joint count at each visit. Indeed, patient self-report measures appear as reproducible as physician-performed formal joint counts [7], in addition to their similar or greater relative efficiencies. Moreover, completion of a self-report questionnaire by a patient in the waiting area provides relevant information to the clinician before seeing the patient.

An international expert committee recommended that a treat-to-target strategy in daily clinical practice should include a validated composite measure of disease activity to guide treatment decisions in routine clinical practice [22]. Collection of RAPID3 scores in the waiting area at each visit ensures that some quantitative data will be available in addition to narrative descriptions of patient status. A careful joint examination, but not necessarily a formal joint count, must also be performed.

Patient completion of an MDHAQ prior to seeing the doctor in no way prevents collection of a formal joint count, DAS28, CDAI or any other measure. As laboratory tests and SJCs are more prognostic of structural damage than patient self-report measures, these measures and a DAS28 or CDAI should be obtained to guide clinical care. Nonetheless, physical function on a patient questionnaire is more significant than radiographs or laboratory tests in the prognosis of most severe outcomes of RA, including work disability, mortality, costs and even joint replacement surgery [23]. A RAPID3 score and DAS28 or CDAI would appear valid for a treat-to-target strategy for optimal care of patients with RA.

> **Rheumatology key messages**
> - Efficacy of treat-to-target in patients with RA appears greater for patient-reported measures than joint counts or laboratory tests.
> - RAPID3 detects differences between RA treatment strategies as effectively as DAS28 and CDAI.
> - RAPID3 may be of value to assess the clinical status of patients with rheumatic diseases in busy clinical settings.

Acknowledgements

Supplementary data

Supplementary data are available at *Rheumatology* Online.

References

1 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.

2 Felson DT. Choosing a core set of disease activity measures for rheumatoid arthritis clinical trials. J Rheumatol 1993;20:531–4.

3 Prevoo MLL, van't Hof MA, Kuper HH et al. 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.

4 Pincus T, Sokka T. Laboratory tests to assess patients with rheumatoid arthritis: advantages and limitations. Rheum Dis Clin North Am 2009;35:731–4.

5 Aletaha D, Nell VPK, Stamm T et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. Arthritis Res Ther 2005;7:R796–806.

6 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.

7 Uhlig T, Kvien TK, Pincus T. Test-retest reliability of disease activity core set measures and indices in rheumatoid arthritis. Ann Rheum Dis 2009;68:972–5.

8 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.

9 Pincus T, Strand V, Koch G 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;48:625–30.

10 Pincus T, Chung C, Segurado OG et al. An index of patient self-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.

11 Pincus T, Bergman MJ, Yazici Y et al. 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 2008;47:345–9.

12 Pincus T, Furer V, Keystone E et al. RAPID3 (routine assessment of patient index data) severity categories and response criteria: similar results to DAS28 and CDAI in the RAPID1 (rheumatoid arthritis prevention of structural damage) clinical trial of certolizumab pegol (CZP). Arthritis Care Res 2011;63:1142–9.

13 Pincus T, Swearingen CJ, Bergman MJ et al. RAPID3 on an MDHAQ is correlated significantly with activity levels of DAS28 and CDAI, but scored in 5 versus more than 90 seconds. Arthritis Care Res 2010;62:181–9.

14 Schipper LG, van Hulst LT, Grol R et al. Meta-analysis of tight control strategies in rheumatoid arthritis: protocolized treatment has additional value with respect to the clinical outcome. Rheumatology 2010;49:2154–64.

15 Soubrier M, Lukas C, Sibilia J et al. Disease activity score-driven therapy versus routine care in patients with recent-onset active rheumatoid arthritis: data from the GUEPARD trial and ESPOIR cohort. Ann Rheum Dis 2011; 70:611–5.

16 Combe B, Benessiano J, Berenbaum F et al. The ESPOIR cohort: a ten-year follow-up of early arthritis in France: methodology and baseline characteristics of the 813 included patients. Joint Bone Spine 2007;74:440–5.

17 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. Arthritis Rheum 2000;43:506–14.

18 Wells G, Li T, Maxwell L et al. 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.

19 Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960;20:37–46.

20 Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33: 159–74.

21 Pincus T, Swearingen CJ, Bergman M et al. 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.

22 Smolen JS, Aletaha D, Bijlsma JW *et al*. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis 2010;69:631–7.

23 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.