



OPEN ACCESS

EXTENDED REPORT

Patient-reported outcomes from a phase 3 study of baricitinib versus placebo or adalimumab in rheumatoid arthritis: secondary analyses from the RA-BEAM study

Edward C Keystone,¹ Peter C Taylor,² Yoshiya Tanaka,³ Carol Gaich,⁴ Amy M DeLozier,⁴ Anna Dudek,⁵ Jorge Velasco Zamora,⁶ Jose Arturo Covarrubias Cobos,⁷ Terence Rooney,⁴ Stephanie de Bono,⁴ Vipin Arora,⁴ Bruno Linetzky,⁴ Michael E Weinblatt⁸

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/annrheumdis-2017-211259>).

¹The Rebecca MacDonald Centre for Arthritis, Mt. Sinai Hospital, Toronto, Canada

²Kennedy Institute of Rheumatology, University of Oxford, Oxford, UK

³University of Occupational and Environmental Health, Fukuoka, Japan

⁴Eli Lilly and Company, Indianapolis, Indiana, USA

⁵Medica Pro Familia, Warsaw, Poland

⁶Instituto Médico CER, Buenos Aires, Argentina

⁷Unidad Reumatologica Las Americas S.C.P. Médico, Mexico, USA

⁸Brigham and Women's Hospital, Boston, Massachusetts, USA

Correspondence to

Professor Edward C Keystone, The Rebecca MacDonald Centre for Arthritis and Autoimmunity, Mt Sinai Hospital, University of Toronto, Toronto, Canada; ed.keystone@sinaihealthsystem.ca

Received 3 February 2017

Revised 22 June 2017

Accepted 24 June 2017

Published Online First

10 August 2017

ABSTRACT

Background To assess the effect of baricitinib on patient-reported outcomes (PROs) in patients with active rheumatoid arthritis and an inadequate response to methotrexate (MTX).

Methods In this double-blind phase 3 study, patients were randomised 3:3:2 to placebo (n=488), baricitinib 4 mg once daily (n=487), or adalimumab 40 mg biweekly (n=330) with background MTX. PROs included the SF-36, EuroQol 5-D (EQ-5D) index scores and visual analogue scale, Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F), Health Assessment Questionnaire-Disability Index (HAQ-DI), Patient's Global Assessment of Disease Activity (PtGA), patient's assessment of pain and Work Productivity and Activity Impairment Questionnaire-Rheumatoid Arthritis (WPAI-RA), and measures collected in electronic patient daily diaries: duration and severity of morning joint stiffness (MJS), Worst Tiredness and Worst Joint Pain. The primary study endpoint was at week 12. Treatment comparisons were assessed with logistic regression for categorical measures or analysis of covariance for continuous variables.

Results Compared with placebo and adalimumab, baricitinib showed statistically significant improvements ($p \leq 0.05$) in HAQ-DI, PtGA, pain, FACIT-F, SF-36 physical component score, EQ-5D index scores and WPAI-RA daily activity at week 12. Improvements were maintained for measures assessed to week 52. Statistically significant improvement in patient diary measures (MJS duration and severity), worst tiredness and worst joint pain were observed for baricitinib versus placebo and adalimumab at week 12 ($p \leq 0.05$).

Conclusions Baricitinib provided significantly greater improvement in most PROs compared with placebo and adalimumab, including physical function MJS, pain, fatigue and quality of life. Improvement was maintained to the end of the study (week 52).

Trial registration NCT01710358.

INTRODUCTION

Rheumatoid arthritis (RA) is characterised by inflammatory activity and joint damage that often result in disability, pain, limitations in physical function and other impairments important to

patients. Outcomes can be improved with effective therapy.¹⁻³ Decreases in physical function may be a consequence of both disease activity and irreversible, progressive joint damage.⁴⁻⁶

Patient-reported outcome (PRO) measures include health-related quality of life (HRQOL), physical function, disability, fatigue, sleep, mental health status, work productivity and work activity impairment.⁷ These are standardised measures, and minimum clinically important differences (MCIDs) have been determined for many. Because these PRO measures are obtained directly from the patients, they may more accurately reflect how the patient feels and functions in relation to RA and to therapy.^{8,9} The PRO measures also may facilitate doctor-patient communication and shared decision making to improve the quality of patient care.¹⁰⁻¹²

Baricitinib is a selective inhibitor of Janus kinase (JAK)1/JAK2 that interrupts signalling in pathways believed to be important in RA pathogenesis. The efficacy of baricitinib has been demonstrated in clinical studies in patients with RA.¹³⁻¹⁷ In the phase 3 RA-BEAM clinical trial (NCT01710358), baricitinib 4 mg once daily (QD) was associated with clinical improvement and inhibition of progression of radiographic joint damage compared with both placebo and adalimumab in patients with RA and an inadequate response to methotrexate (MTX). Specifically, 70% of patients treated with baricitinib achieved the American College of Rheumatology 20% response rate compared with 40% of placebo-treated patients and 61% of adalimumab-treated patients. This manuscript describes the PRO data collected in the RA-BEAM clinical trial of baricitinib.¹⁸

METHODS

Patients and study design

RA-BEAM was a randomised, double-blind, double-dummy, placebo-controlled and active-controlled, parallel-arm, 52-week study conducted at 281 centres in 26 countries. Detailed methods of the RA-BEAM study have been published previously.¹⁸ Briefly, patients were ≥ 18 years old with active RA ($\geq 6/68$ tender and $\geq 6/66$ swollen joints; serum



CrossMark

To cite: Keystone EC, Taylor PC, Tanaka Y, et al. *Ann Rheum Dis* 2017;**76**:1853-1861.

high-sensitivity C reactive protein ≥ 6 mg/L). Patients had inadequate response to MTX. At baseline, patients were required to have either ≥ 3 joint erosions (based on radiographs) or > 1 joint erosion with seropositivity for rheumatoid factor or anticitrullinated peptide antibodies.

Patients were randomised 3:3:2 to receive placebo, baricitinib 4 mg once daily, or biweekly subcutaneous adalimumab 40 mg, in addition to their existing background therapy (including MTX). The primary analysis time point for the study was at week 12. At week 24, patients receiving placebo switched to baricitinib. At week 16, patients whose tender and swollen joint counts improved from baseline by $< 20\%$ at both weeks 14 and 16 were assigned rescue treatment (baricitinib 4 mg). After week 16, patients could be rescued at investigators' discretion based on joint counts. The study was conducted in accordance with ethical principles of the Declaration of Helsinki and Good Clinical Practice guidelines and was approved by each centre's institutional review board or ethics committee. All patients provided written informed consent.

Study endpoints and assessments

PROs were prespecified as secondary objectives of the study. Physical function was measured using the Health Assessment Questionnaire-Disability Index (HAQ-DI).^{19, 20} Scores range from 0 to 3, with lower scores reflecting better physical function and, thus, less disability. The HAQ-DI score changes were assessed in the context of an MCID of 0.22,²¹ and the percentage of patients who reported scores that met or exceeded the population normative value (< 0.5) was also assessed.²² Disease activity and arthritis pain were measured using the Patient's Global Assessment of Disease Activity (PtGA) and the patient's assessment of pain visual analogue scales (VAS; 0–100 mm) in which higher scores indicate more disease activity or pain. Fatigue was assessed by the Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) scale (range 0–52), with higher scores representing less fatigue.²³ A 3-point to 4-point change has been considered an MCID,^{23–25} and in this study a value of 3.56²⁵ was used to assess the clinical relevance of changes in FACIT-F scores. The percentage of patients who reported scores that met or exceeded the population normative value (≥ 40.1) was also assessed.²³

Some PROs were recorded using a daily electronic diary (referred to as diary PROs) from day 1 through week 12. These included duration of morning joint stiffness (MJS), and the novel MJS Severity, Worst Tiredness and Worst Joint Pain numeric rating scales (NRS). The scores for the NRS range from 0 to 10, with 10 being the worst level.

HRQOL was evaluated using the Medical Outcomes Study Short-Form-36 (SF-36; version 2, Acute),^{26, 27} which assesses eight domains scored from 0 to 100 that are normalised into physical component score (PCS) and mental component score (MCS). An MCID of 5 was used to assess the clinical relevance of changes in SF-36 scores.^{28, 29} A sensitivity analysis with an MCID of 2.5 was also evaluated. The EuroQoL 5-Dimensions (EQ-5D) Health State Profile was also used to assess HRQOL. The EQ-5D consists of two components: a descriptive system of the respondent's health and a rating of their current health state (0–100 mm VAS).³⁰ The UK and US scoring algorithms provide an index score using the UK or US population weighting to normalise it to that population.^{31, 32}

The Work Productivity and Activity Impairment Questionnaire-Rheumatoid Arthritis (WPAI-RA) was used to measure the health and symptoms of overall work productivity and

impairment of regular activities, as measured during the past 7 days. Scores are calculated as impairment percentages³³ with higher percentages indicating greater impairment and less productivity.

Non-diary PROs were assessed at baseline and at weeks 1, 2, 4 and every 4 weeks thereafter to week 32 and after week 32, they were assessed at weeks 40 and 52. There were, however, some exceptions to the schedule before week 4. The FACIT-F, SF-36 and the EQ-5D were assessed at baseline and week 4, and the WPAI-RA was assessed at baseline, week 2 and week 4. These measures followed the same schedules as the other PROs after Week 4.

Statistical analyses

Randomised patients treated with ≥ 1 dose of placebo, baricitinib or adalimumab were included in the efficacy analyses on the basis of a modified intention-to-treat principle (analysis set).

Least squares mean (LSM) change from baseline for treatment comparisons of continuous efficacy variables were obtained using analysis of covariance with treatment, geographical region, baseline joint erosion status and baseline value in the model. For diary PRO data, analyses were based on the average of scores collected in the 7 days prior to the study visit, without baseline adjustment, until the week 12 visit date; daily scores from the day of randomisation (day 1) up to day 28 were also assessed. For the daily score analysis, mixed models for repeated measures were applied, with duration of MJS analysed by non-parametric methods.

As per the predefined analysis plan, patients who were rescued or discontinued from study or study treatment were thereafter defined as non-responders (non-responder imputation) for analysis of all categorical efficacy measures. For continuous efficacy measures, modified last observation carried forward was used, where the last observation before rescue or discontinuation was used for all subsequent time points. The WPAI-RA measures were censored after rescue or discontinuation without imputation applied.

All analyses are based on a significance level of 0.05 (two sided). *p* Values were not adjusted for multiple comparisons.

RESULTS

Patient disposition and baseline characteristics

A total of 1307 patients were randomised (488 placebo, 487 baricitinib 4 mg and 330 adalimumab) and 1305 patients received treatment. Patient disposition has been previously reported.¹⁸ Baseline patient characteristics and disease activity were similar among groups (online supplementary file 1).¹⁸ Most patients ($> 99\%$) were receiving background MTX. The majority of patients had received ≥ 2 prior conventional synthetic disease modifying antirheumatic drugs (csDMARDs). Baseline PROs indicated a significant disease burden, consistent with the baseline clinical disease activity (online supplementary file 1).

Patient-reported outcomes

HAQ-DI, PtGA and pain

As reported in Taylor, *et al*,¹⁸ for HAQ-DI, PtGA, and the patient's assessment of pain, statistically significant improvements in the baricitinib group versus placebo were evident as early as week 1, the first assessment. Significant improvements in physical function and reductions in PtGA and pain were maintained at week 12 and through week 52, the end of the study (table 1). When compared with adalimumab, statistically significant improvements in HAQ-DI were seen as early as week 4 and

Table 1 Least squares mean change from baseline at 12 and 52 weeks for PRO

PRO measures (95% CI)	Week 12			Week 52	
	Placebo (n=488)	Baricitinib (n=487)	Adalimumab (n=330)	Baricitinib (n=487)	Adalimumab (n=330)
Physical function (HAQ-DI)	-0.34 (-0.39, -0.29)	-0.66*** †† (-0.71, -0.61)	-0.56*** (-0.62, -0.50)	-0.77†† (-0.83, -0.71)	-0.66 (-0.73, -0.59)
Patient's Global Assessment of Disease Activity (PtGA)	-16.7 (-18.9, -14.6)	-31.2*** †† (-33.3, -29.1)	-26.6*** (-29.1, -24.1)	-36.3††† (-38.7, -33.9)	-30.3 (-33.1, -27.5)
Patient's Assessment of Pain EuroQol-5-Dimensions (EQ-5D)	-17.1 (-19.4, -14.9)	-31.5*** †† (-33.7, -29.3)	-26.4*** (-29.0, -23.7)	-36.1††† (-38.6, -33.7)	-30.3 (-33.1, -27.5)
Health State Index Score, UK algorithm	0.102 (0.084, 0.119)	0.184*** (0.167, 0.202)	0.167*** (0.146, 0.188)	0.217† (0.197, 0.238)	0.182 (0.158, 0.206)
Health State Index Score, US algorithm	0.071 (0.058, 0.083)	0.130*** (0.118, 0.142)	0.117*** (0.102, 0.131)	0.154† (0.139, 0.169)	0.129 (0.112, 0.146)
VAS	7.7 (5.6, 9.8)	14.8*** †† (12.8, 16.9)	10.1 (7.7, 12.6)	19.1††† (16.6, 21.5)	11.6 (8.8, 14.5)

*p≤0.05, **p≤0.01, ***p≤0.001 versus placebo.

†p≤0.05, ††p≤0.01, †††p≤0.001 versus adalimumab.

HAQ-DI, Health Assessment Questionnaire-Disability Index; PRO, patient-reported outcomes; VAS, visual analogue scale.

at week 2 for PtGA and pain, respectively; these improvements were maintained at week 12 and through week 52.

The percentages of patients who reported improvements that met or exceeded the HAQ-DI MCID of ≥0.22 at week 12 for placebo, baricitinib and adalimumab, respectively, were 58%, 75%, and 71% (p≤0.001 for baricitinib vs placebo and adalimumab vs placebo; p=0.302 for baricitinib vs adalimumab) and were 68% and 58% at week 52 for baricitinib vs adalimumab (p≤0.01). The percentage of patients who reported scores that met or exceeded the population normative value of <0.5 at week 12 or those who met or exceeded at week 52 ranged from 24% to 32% for baricitinib and adalimumab (online supplementary file 2). The percentages for baricitinib and adalimumab were statistically different (p<0.05) than placebo at week 12.

MJS duration, MJS severity, worst tiredness and worst joint pain

Baricitinib treatment resulted in significant improvement versus placebo and adalimumab for all four measures at the primary time point of the study, week 12 (table 2). Improvements versus placebo were significant from week 1 for severity of MJS, Worst Tiredness and Worst Joint Pain and from week 2 for the duration

of MJS. Improvements of baricitinib versus adalimumab were observed from as early as week 2 for Worst Joint Pain, week 4 for severity of MJS and week 8 for Worst Tiredness; for duration of MJS, improvements were statistically different at weeks 1 and 12, as reported by Taylor *et al.*¹⁸

Consistent with the weekly averaged data, the daily diary scores showed significant improvement in patients receiving baricitinib compared with both placebo and adalimumab. Improvements relative to placebo were observed as early as day three for the severity of MJS, Worst Tiredness, and Worst Joint Pain and by day five for the duration of MJS (figure 1). Improvements relative to adalimumab were observed as early as day 19 for the severity of MJS, day 21 for Worst Tiredness and day 17 for Worst Joint Pain. No significant differences were observed in the first 28 days between baricitinib and adalimumab for the duration of MJS.

FACIT-F

Treatment with baricitinib or adalimumab was associated with significant improvements in FACIT-F at the first assessment of the measure at week 4 (p≤0.001 for baricitinib

Table 2 Day 1, week 1 and week 12 data from patient daily diaries

PRO measures	Day 1 median (IQR)			Week 1 median (IQR)			Week 12 median (IQR)		
	Placebo (n=488)	Baricitinib (n=487)	Adalimumab (n=330)	Placebo (n=488)	Baricitinib (n=487)	Adalimumab (n=330)	Placebo (n=488)	Baricitinib (n=487)	Adalimumab (n=330)
Duration of morning joint stiffness, minutes, median (IQR)	60.0 (30.0, 180.0)	60.0 (30.0, 180.0)	60.0 (20.0, 180.0)	87.5 (32.5, 180.0)	75.0† (27.5, 154.3)	60.0*** (20.0, 150.0)	60.0 (17.1, 154.3)	27.1*** † (4.3, 90.0)	36.6*** (9.2, 120.0)
	Day 1 (mean (SD))			Week 1 LSM (95% CI)			Week 12 LSM (95% CI)		
Severity of morning joint stiffness	5.5 (2.2)	5.4 (2.2)	5.3 (2.3)	5.3 (5.1 to 5.4)	4.8*** (4.6 to 4.9)	4.7*** (4.5 to 4.8)	4.1 (3.9 to 4.3)	3.0*** †† (2.8 to 3.2)	3.5*** (3.2 to 3.7)
Worst Tiredness	5.6 (2.2)	5.6 (2.2)	5.5 (2.2)	5.3 (5.1 to 5.4)	4.9*** (4.7 to 5.0)	4.8*** (4.7 to 5.0)	4.3 (4.1 to 4.5)	3.6*** † (3.4 to 3.8)	3.9** (3.6 to 4.1)
Worst Joint Pain	5.9 (2.1)	5.9 (2.1)	5.7 (2.2)	5.6 (5.5 to 5.8)	5.0*** (4.9 to 5.2)	5.1*** (4.9 to 5.2)	4.6 (4.4 to 4.8)	3.4*** ††† (3.2 to 3.6)	4.0*** (3.8 to 4.2)

*p≤0.05; **p≤0.01; ***p≤0.001 versus placebo.

†p≤0.05, ††p≤0.01; †††p≤0.001 versus adalimumab.

LSM, least squares mean; PRO, patient-reported outcome.

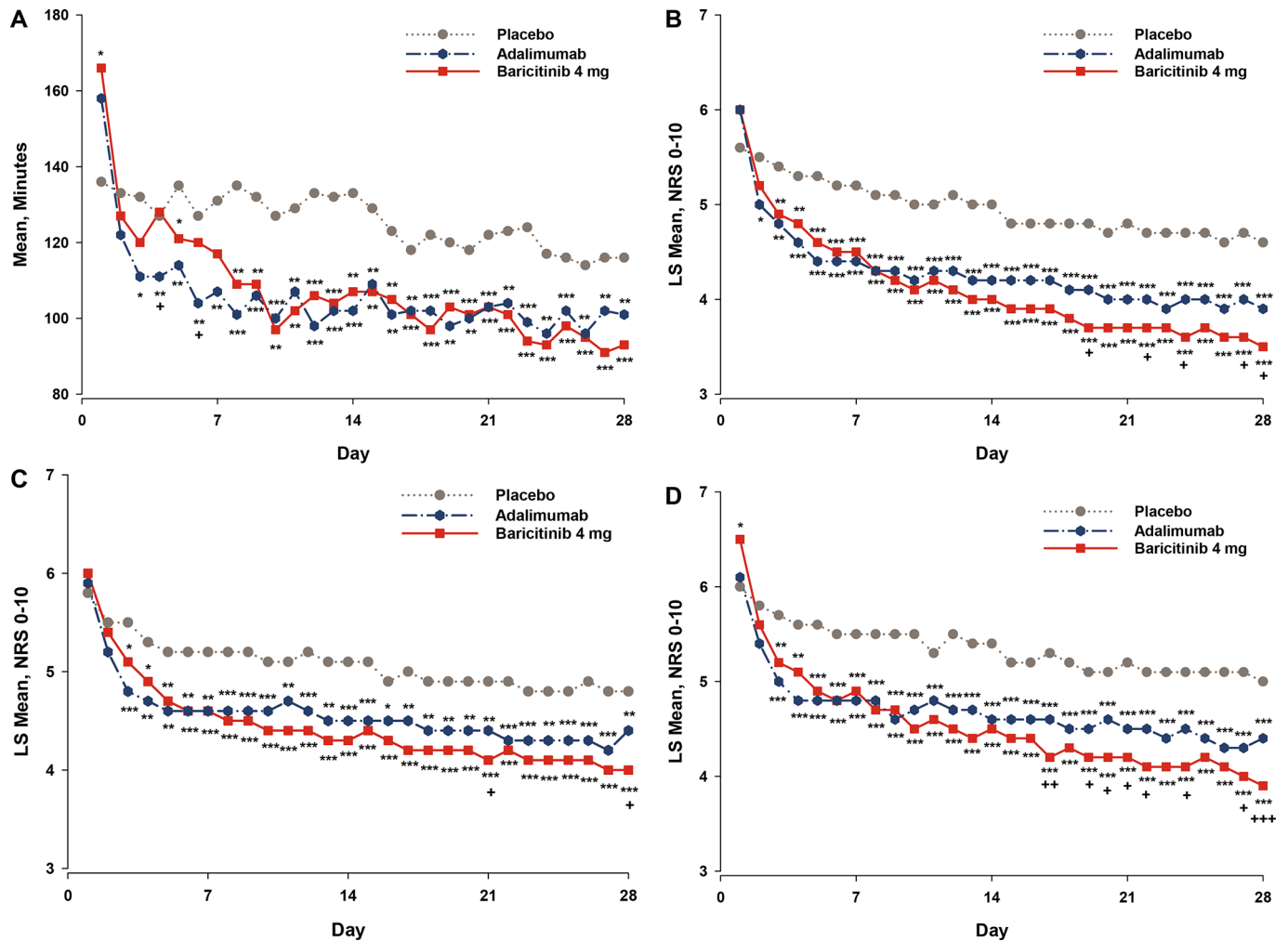


Figure 1 Change from baseline over time for the patient-reported outcomes collected by the daily patient electronic diaries (data and either SD or 95% CIs are presented in online supplementary file 5). (A) Duration of Morning Joint Stiffness: data are mean duration of morning joint stiffness in minutes, based on daily diary entries. Daily question: 'Please indicate how long your morning joint stiffness lasted today'. Indications of statistical significance based on analysis of median difference. (B) Severity of Morning Joint Stiffness: data are LS mean scores for severity of morning joint stiffness, based on daily diary entries. Higher values indicate greater severity with numeric rating scale anchors (0–10). Daily question: 'Please rate the overall level of morning joint stiffness you had from the time you woke up today'. (C) Worst Tiredness: data are LS mean scores for Worst Tiredness, based on daily diary entries. Higher values indicate greater tiredness with numeric rating scale anchors (0–10). Daily question: 'Please rate your tiredness by selecting the one number that describes your tiredness at its worst in the last 24 hours'. (D) Worst Joint Pain: data are LS mean scores for Worst Joint Pain, based on daily diary entries. Higher values indicate greater pain with numeric rating scale anchors (0–10). Daily question: 'Please rate your joint pain by selecting the one number that describes your joint pain at its worst in the last 24 hours'. LS, least squares; NRS, numeric rating scale. p Value versus placebo: * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$. p Value versus adalimumab: + $p \leq 0.05$; ++ $p \leq 0.01$; +++ $p \leq 0.001$.

vs placebo; $p \leq 0.01$ for adalimumab vs placebo; figure 2). The improvements in the FACIT-F score were sustained to week 24 for both baricitinib and adalimumab versus placebo ($p \leq 0.001$) and were significant at weeks 20, 28 and 52 for baricitinib versus adalimumab ($p \leq 0.05$; figure 2).

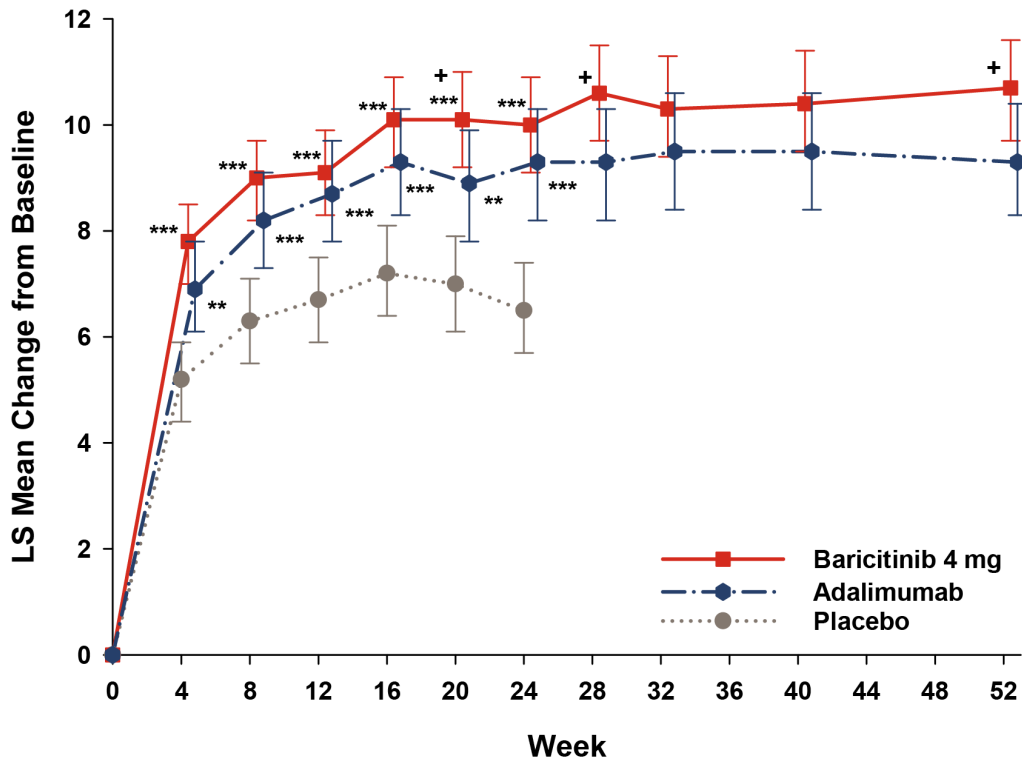
For the FACIT-F, the percentage of patients who reported improvements that met or exceeded the MCID (≥ 3.56) was 59%, 66% and 68% for placebo, baricitinib and adalimumab, respectively ($p \leq 0.05$ for baricitinib vs placebo; $p \leq 0.01$ for adalimumab vs placebo), at week 12 and were 60% and 54% at week 52 for baricitinib and adalimumab, respectively ($p = 0.084$; figure 2). The percentage of patients who reported scores that met or exceeded the population normative value of ≥ 40.1 ranged from 41% to 46% for baricitinib and adalimumab at weeks 12 and 52 (online supplementary file 2).

Health-related quality of life

SF-36

Patients treated with baricitinib or adalimumab showed statistically significantly improved differences compared with placebo in most of the eight SF-36 domains at week 12 except for the mental health domain (both baricitinib and adalimumab) and role-emotional domain (adalimumab), which improved but did not achieve statistical significance (table 3). Compared with adalimumab, patients treated with baricitinib showed statistically significant improvement in most of the domains at week 52, except for the mental health domain (table 3).

Compared with placebo, SF-36 PCS was statistically significantly improved for patients treated with baricitinib and adalimumab (figure 3A) from the first postbaseline assessment at week 4 and was maintained through weeks 12 and 52. At week 12, the percentage of patients who met or exceeded the MCID (≥ 5) for



% of patients who met or exceeded the MCID ≥ 3.56										
Week	4	8	12	16	20	24	28	32	40	52
Placebo (N=488)	53	56	59	56	52	43	---	---	---	---
Baricitinib 4 mg (N=487)	63**	67***	66*	69***	68***	65***	64+	62	62+	60
Adalimumab (N=330)	59	66**	68**	65*	62**	59***	57	57	55	54

Figure 2 Change from baseline over time for the FACIT-F. Higher scores indicate less fatigue. Range=0–52. FACIT-F, Functional Assessment of Chronic Illness Therapy-Fatigue; MCID, minimum clinically important differences. p Value versus placebo: *p≤0.05; **p≤0.01; ***p≤0.001. p Value versus adalimumab: +p≤0.05; ++p≤0.01; +++p≤0.001.

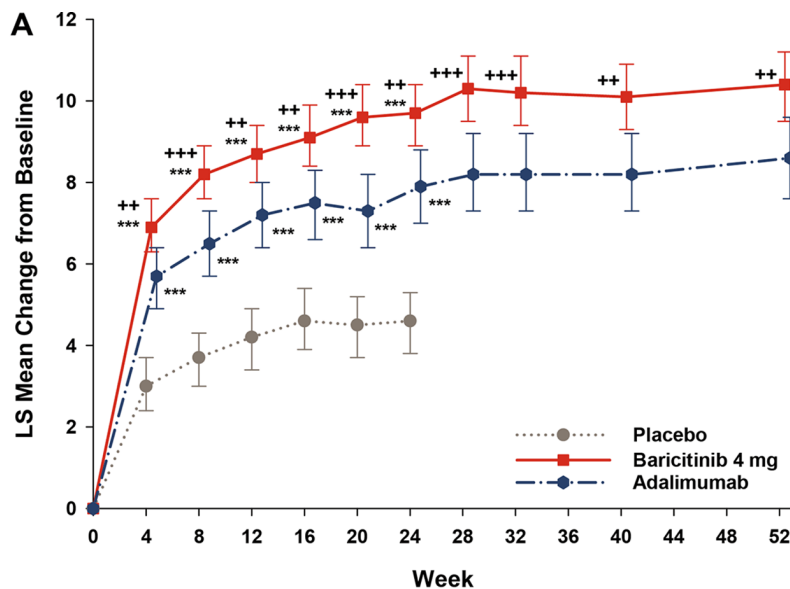
placebo, baricitinib and adalimumab, respectively, was 40%, 65% and 56% at week 12 (for both groups vs placebo, p≤0.001 and baricitinib vs adalimumab, p≤0.05) and was 60% and 52% at week 52 for baricitinib versus adalimumab (p≤0.05). For the

SF-36 MCS measure, numeric, but not statistically significant differences in the change from baseline were found for both baricitinib and adalimumab versus placebo at all time points, except for baricitinib versus placebo at week 24 (p≤0.01). The

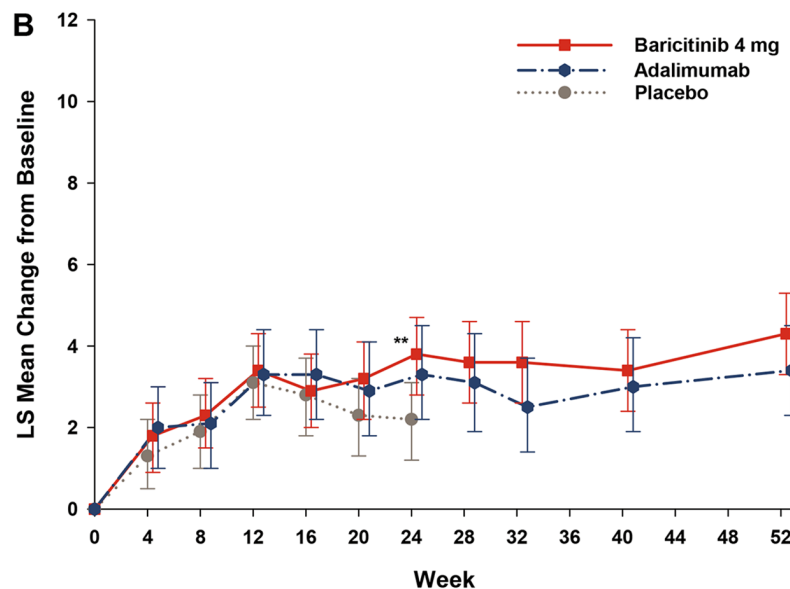
Table 3 Baseline values and least squares mean changes from baseline at weeks 12 and 52 for SF-36 domain scores.

SF-36 domain scores	Placebo (n=488)	Baseline mean (SD)		Placebo (n=488)	Week 12 LSM		Week 52 LSM	
		Baricitinib (n=487)	Adalimumab (n=330)		Baricitinib (n=487)	Adalimumab (n=330)	Baricitinib (n=487)	Adalimumab (n=330)
Physical functioning	32.4 (10.4)	32.3 (10.2)	31.6 (10.7)	4.3	8.0***†	6.8***	9.9†	8.4
Role physical	36.3 (10.3)	35.5 (10.3)	34.5 (10.5)	4.4	7.8***	6.7***	9.4††	7.5
Bodily pain	34.9 (7.7)	34.6 (7.5)	34.5 (8.5)	4.6	9.1***†	7.6***	11.2†	9.7
General health	36.6 (8.6)	37.3 (8.1)	36.3 (8.7)	3.1	5.4***	4.5†††	6.1†	4.8
Vitality	43.9 (10.1)	43.8 (9.5)	43.2 (10.5)	3.9	6.4***	5.7†††	7.9†	6.6
Social functioning	41.3 (11.3)	40.9 (11.6)	40.0 (12.2)	3.0	5.6***	4.4*	6.6††	4.6
Role emotional	41.4 (12.5)	41.4 (12.5)	40.3 (12.9)	3.7	5.1*	4.8	6.6†	5.3
Mental health	42.9 (11.3)	43.3 (11.1)	42.5 (11.5)	3.7	4.0	3.9	5.1	4.4

*p≤0.05, **p≤0.01, ***p≤0.001 versus placebo.
 †p≤0.05, ††p≤0.01, †††p≤0.001 versus adalimumab.
 LSM, least squares mean; SF-36, Short-Form-36.



% of patients who met or exceeded the MCID ≥ 5										
Week	4	8	12	16	20	24	28	32	40	52
Placebo (N=488)	34	39	40	41	40	35	---	---	---	---
Baricitinib 4 mg (N=487)	56***	64****	65****	67****	68****	65****	66***	61*	59	60*
Adalimumab (N=330)	51***	54***	56***	56***	55***	57***	54	53	53	52



% of patients who met or exceeded the MCID ≥ 5										
Week	4	8	12	16	20	24	28	32	40	52
Placebo (N=488)	31	32	33	32	27	28	---	---	---	---
Baricitinib 4 mg (N=487)	33	33	37	36	36**	38**	34	36	32	34
Adalimumab (N=330)	32	34	37	36	34*	34	33	32	32	29

Figure 3 Change from baseline for the physical and mental component score for the SF-36. (A) Physical component score: data in table are % patients who met or exceeded the minimum clinically important difference in SF-36 PCS (≥ 5 points). Higher scores indicate improvement. (B) Mental component score: data in table are % patients who met or exceeded the minimum clinically important difference in SF-36 MCS (≥ 5 points). Higher scores indicate improvement. p Value versus placebo: * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$. p Value versus adalimumab: + $p \leq 0.05$; ++ $p \leq 0.01$; +++ $p \leq 0.001$.

proportion of patients who reported improvements that met or exceeded the MCID for the MCS was not statistically significantly different from placebo for either group at any time point, except at weeks 20 (both baricitinib and adalimumab differed from placebo, $p \leq 0.05$) and at week 24 ($p \leq 0.01$ for baricitinib vs placebo; figure 3B). Results were similar for the MCID value of 2.5 (online supplementary file 3).

EQ-5D

A statistically significant improvement in the EQ-5D index scores (both US and UK) were observed at the first postbaseline assessment, week 4 (data not shown), for both baricitinib and adalimumab versus placebo and was maintained to week 12 (table 1). By week 52, statistically significant improvements in EQ-5D index scores were observed for baricitinib versus adalimumab (table 1). For the EQ-5D VAS at 4 weeks, baricitinib-treated and adalimumab-treated patients showed statistically significant improvement compared with placebo-treated patients (data not shown). By week 12, however, statistically significant improvement in EQ-5D VAS was observed for only baricitinib-treated patients ($p \leq 0.001$ vs placebo; $p \leq 0.01$ versus adalimumab); this was maintained through week 52 for baricitinib versus adalimumab ($p \leq 0.001$; table 1).

Work Productivity and Activity Impairment

At baseline, 41%–43% of the patients were employed. Patients treated with baricitinib reported statistically significantly improved daily activity compared with placebo and adalimumab at week 12 ($p \leq 0.001$ for both groups vs placebo; $p \leq 0.01$ for baricitinib vs adalimumab); improvements compared with adalimumab, however, were not statistically significant at week 52 (online supplementary file 4). Among those patients who were employed at baseline and those who maintained employment at week 12, statistically significant improvements in absenteeism ($p \leq 0.05$), presenteeism ($p \leq 0.001$) and work productivity loss ($p \leq 0.001$) were seen with baricitinib compared with placebo; improvements compared with adalimumab, however, were not statistically significant at week 52. Only work productivity loss was statistically significantly improved with baricitinib versus adalimumab at week 12 ($p \leq 0.05$; online supplementary file 3).

DISCUSSION

The RA-BEAM study evaluated baricitinib 4 mg once daily in patients with an inadequate response to MTX who were naive to biological DMARDs using placebo and adalimumab 40 mg biweekly as comparators.¹⁸ Patients continued to take stable background csDMARDs (including MTX) during the study. This paper evaluates whether the clinical efficacy data for baricitinib were complemented by corresponding changes in PROs.

Baseline PROs describe substantial duration (≥ 60 min) and severity of MJS, severe impairment of physical function and high levels of pain and fatigue (including tiredness) among patients enrolled in the study. Baricitinib treatment produced significantly greater improvements compared with placebo and adalimumab in most of the prespecified PROs, including physical function, pain, fatigue, duration and severity of MJS and HRQOL at week 12. Furthermore, baricitinib produced rapid improvements in the diary PROs compared with placebo and adalimumab, with significant differences vs placebo appearing within days of initiating treatment. Improvements were maintained to week 52 compared with adalimumab in physical function, pain, fatigue and HRQOL (eg, SF-36 PCS and EQ-5D).

In this analysis and in the results presented by Taylor *et al*,¹⁸ treatment with baricitinib resulted in a rapid improvement in PROs; patients showed statistically significant improvements as early as week 1 in HAQ-DI, PtGA and the patient's assessment of pain, and these results were maintained until the end of the trial at week 52. Compared with placebo, a significantly greater proportion of patients treated with baricitinib or adalimumab reported improvements that met or exceeded the MCID and the population normative values for HAQ-DI and FACIT-F at week 12.

Similar results were seen for duration and severity of MJS, Worst Tiredness and Worst Joint Pain as assessed using the patient daily diaries and improvement continued to week 12. The rapid onset of action, with improvement in relevant signs and symptoms (such as pain and tiredness) as early as 3 days after the start of treatment, is a useful complement to the efficacy observed at the later time points.

Consistent with these results, patients treated with baricitinib reported improvements in HRQOL, as measured by the EQ-5D and SF-36 PCS compared with placebo and adalimumab. For the SF-36, improvements across most of the SF-36 domains were observed for baricitinib and adalimumab compared with placebo at week 12 and for baricitinib compared with adalimumab at week 52. Furthermore, when a five-point change was used for the MCID, it was found that 65% of the baricitinib-treated patients met or exceeded the MCID for the SF-36 PCS (figure 3A). In contrast with the PCS, no statistically significant differences were observed between baricitinib-treated and adalimumab-treated patients compared with placebo-treated patients with the SF-36 MCS. Across treatment groups at baseline, the SF-36 MCS values ranged from 46 to 47, which are close to the population normative data of 50.²⁸ This suggests only modest impairment for the MCS at baseline such that a marked improvement with therapy would not be expected. This SF-36 MCS result aligns with previous results from other trials.^{34–36}

Compared with placebo, patients treated with baricitinib showed statistically significant improvement across all scores of the work productivity assessment at week 12. When compared with adalimumab, the baricitinib-treated patients showed statistically significant improvement in work productivity loss and impairment of regular activity at week 12; these improvements continued through week 52 but were not statistically significantly different.

The results from this analysis are similar to those observed in other phase 3, randomised clinical trials of baricitinib in different patient populations.^{34–36}

The limitations of this analysis include the use of carrying forward the last observations before rescue or discontinuation. This method assumes that the PRO values do not change over time. Also, as in most double-blind comparator trials, the inclusion and exclusion criteria restrict patient participation such that these results may not be fully generalisable to the population seen in clinical practice.

This study used well-established PRO measures that can holistically evaluate the burden of RA and the treatment effects across many health domains. Some of the PRO measures are incorporated into the patient ratings in the ACR core set, while others, such as the EQ-5D and SF-36, are established HRQOL instruments that may more broadly measure the effects of RA and treatment on patients. The use of diary records allowed patients to report the impact of symptoms of importance to them as they arose, therefore permitting a more complete evaluation than by means of periodic recording of recollected symptoms. Furthermore,

these PRO measures may help facilitate discussions between patients and their healthcare providers; they may help address patient concerns such as how long it will take to feel improvement (onset of action), how long to try the new treatment before determining that it is not effective (efficacy plateau) and how long the treatment will be effective (sustainability or the risk of relapse). In addition to facilitating the physician–patient dynamic, PRO measures are being increasingly used in randomised clinical trials and allow for epidemiological assessments across different patient populations and disease states. PRO assessments such as the work productivity measure also provide an insight into the broader, societal impact of RA.

The use of a variety of PRO measures also allows for an assessment of the clinical importance of the present study's results. Similar trends were observed between comparisons of baricitinib with both placebo and adalimumab in many of the PROs. Additionally, some PRO measures were assessed with established and validated MCID values. Statistically significant differences in MCIDs between treatment groups implies clinical significance on a group level. Collectively, the results of the present study demonstrate treatment benefits for baricitinib that appear clinically relevant.

The RA-BEAM study demonstrated that patients treated with baricitinib experienced a greater improvement compared with patients treated with placebo or adalimumab in most PROs across different domains of RA, including physical function, MJS, fatigue, pain and HRQOL. These improvements tended to occur within the first weeks of treatment and were maintained throughout the 52-week trial.

Correction notice This article has been corrected since it published Online First. The 'patient disposition and baseline characteristics' paragraph has been updated.

Acknowledgements The authors would like to thank Barbara Coffey, MS, and Molly Tomlin, MS, of Eli Lilly and Company for their assistance with manuscript preparation and process support and Scott Beattie, PhD, for statistical support.

Contributors All authors have contributed to the work and approve the presented findings.

Funding This study was funded by Eli Lilly and Company and Incyte Corporation.

Competing interests ECK has received grant/research support, consulting support or speaker bureau fees from Abbott, Amgen, AstraZeneca, Biotest, Bristol Myers Squibb, F. Hoffmann-La Roche Inc., Janssen, Eli Lilly and Company, Genentech, Janssen, Merck, Novartis, Pfizer, Sanofi-Aventis and UCB. PCT has received grant/research support or consulting support from AbbVie, Bristol Myers Squibb, Celgene, Eli Lilly and Company, Galapagos, GlaxoSmithKline, Merck, Pfizer, Takeda and UCB. YT has received grant/research support or speaker bureau fees from AbbVie, Astellas, Asahi Kasei, Bristol Myers Squibb, Chugai, Daiichi-Sankyo, Eisai, Eli Lilly and Company, GlaxoSmithKline, Janssen, Kyowa-Kirin, Mitsubishi-Tanabe, MSD, Pfizer, Santen, Takeda, Taisho-Toyama, Teijin and UCB. CG, AMD, TR, SdB, VA and BL are employees of Eli Lilly and Company, and all except AMD are shareholders of Eli Lilly and Company. AD is a consultant for Eli Lilly and Company. JVZ is a consultant for Eli Lilly and Company. JACC is a consultant for Eli Lilly and Company. MEW has received grant/research support or consulting support from AbbVie, Amgen, Bristol-Myers Squibb, Canfit, Corrona, Crescendo Bioscience, Genzyme, Idera, Infinity, Lycera, Eli Lilly and Company, Medimmune, Merck, Merck/serona, Momenta, Novartis, Pfizer, Roche, Samsung, Sanofi and UCB.

Ethics approval The study was approved by each center's institutional review board or ethics committee.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

- Carr A, Hewlett S, Hughes R, *et al.* Rheumatology outcomes: the patient's perspective. *J Rheumatol* 2003;30:880–3.
- Strand V, Cohen S, Crawford B, *et al.* Patient-reported outcomes better discriminate active treatment from placebo in randomized controlled trials in rheumatoid arthritis. *Rheumatology* 2004;43:640–7.
- Klarenbeek NB, Güler-Yüksel M, van der Kooij SM, *et al.* The impact of four dynamic, goal-steered treatment strategies on the 5-year outcomes of rheumatoid arthritis patients in the BeSt study. *Ann Rheum Dis* 2011;70:1039–46.
- Drossaers-Bakker KW, de Buck M, van Zeben D, *et al.* Long-term course and outcome of functional capacity in rheumatoid arthritis: the effect of disease activity and radiologic damage over time. *Arthritis Rheum* 1999;42:1854–60.
- Aletaha D, Smolen J, Ward MM. Measuring function in rheumatoid arthritis: identifying reversible and irreversible components. *Arthritis Rheum* 2006;54:2784–92.
- Smolen JS, Aletaha D, Grisar JC, *et al.* Estimation of a numerical value for joint damage-related physical disability in rheumatoid arthritis clinical trials. *Ann Rheum Dis* 2010;69:1058–64.
- Kekow J, Moots RJ, Emery P, *et al.* Patient-reported outcomes improve with etanercept plus methotrexate in active early rheumatoid arthritis and the improvement is strongly associated with remission: the COMET trial. *Ann Rheum Dis* 2010;69:222–5.
- Gossec L, Dougados M, Dixon W. Patient-reported outcomes as end points in clinical trials in rheumatoid arthritis. *RMD Open* 2015;1:e000019.
- Guidance for industry: patient-reported outcome measures: use in Medical Product Development to support labeling claims. Secondary Guidance for Industry: Patient-Reported Outcome Measures Use in Medical Product Development to Support Labeling Claims. 2009 <http://www.fda.gov/downloads/Drugs/UGuidances/UCM193282.pdf> (accessed 17 Apr 2017).
- Fries JF, Spitz P, Kraines RG, *et al.* Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23:137–45.
- Sokka T. Morning stiffness and other patient-reported outcomes of rheumatoid arthritis in clinical practice. *Scand J Rheumatol Suppl* 2011;125:23–7.
- Smolen JS, Landewe R, Bijlsma J, *et al.* EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017;1–18.
- Keystone EC, Taylor PC, Drescher E, *et al.* Safety and efficacy of baricitinib at 24 weeks in patients with rheumatoid arthritis who have had an inadequate response to methotrexate. *Ann Rheum Dis* 2015;74:333–40.
- Tanaka Y, Emoto K, Cai Z, *et al.* Efficacy and safety of Baricitinib in Japanese patients with active rheumatoid Arthritis Receiving Background Methotrexate therapy: a 12-week, Double-blind, randomized Placebo-controlled study. *J Rheumatol* 2016;43:504–11.
- Fleischmann R, Schiff M, van der Heijde D, *et al.* Baricitinib, Methotrexate, or combination in patients with Rheumatoid Arthritis and no or Limited Prior Disease-Modifying Antirheumatic Drug Treatment. *Arthritis Rheumatol* 2017;69:506–17.
- Dougados M, van der Heijde D, Chen YC, *et al.* Baricitinib in patients with inadequate response or intolerance to conventional synthetic DMARDs: results from the RA-BUILD study. *Ann Rheum Dis* 2017;76:88–95.
- Genovese MC, Kremer J, Zamani O, *et al.* Baricitinib in patients with refractory rheumatoid Arthritis. *N Engl J Med* 2016;374:1243–52.
- Taylor PC, Keystone EC, van der Heijde D, *et al.* Baricitinib versus Placebo or Adalimumab in Rheumatoid Arthritis. *N Engl J Med* 2017;376:652–62.
- Bruce B, Fries JF. The Health Assessment Questionnaire (HAQ). *Clin Exp Rheumatol* 2005;23(Suppl 39):S14–18.
- Ramey D, Fries J, Singh G. The Health Assessment Questionnaire 1995: status and review. Spiker B, ed. *Quality of life and pharmacoconomics in clinical trials*. 2nd ed. Philadelphia: Lippincott-Raven, 1996:227–37.
- Wells GA, Tugwell P, Kraag GR, *et al.* Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. *J Rheumatol* 1993;20:557–60.
- Navarro-Compán V, Smolen JS, Huizinga TW, *et al.* Quality indicators in rheumatoid arthritis: results from the METEOR database. *Rheumatology* 2015;54:1630–9.
- Cella D, Yount S, Sorensen M, *et al.* Validation of the Functional Assessment of Chronic Illness Therapy Fatigue Scale relative to other instrumentation in patients with rheumatoid arthritis. *J Rheumatol* 2005;32:811–9.
- Strand V, Burmester GR, Zerbini CA, *et al.* Tofacitinib with methotrexate in third-line treatment of patients with active rheumatoid arthritis: patient-reported outcomes from a phase III trial. *Arthritis Care Res* 2015;67:475–83.
- Keystone E, Burmester GR, Furie R, *et al.* Improvement in patient-reported outcomes in a rituximab trial in patients with severe rheumatoid arthritis refractory to anti-tumor necrosis factor therapy. *Arthritis Rheum* 2008;59:785–93.
- Brazier JE, Harper R, Jones NM, *et al.* Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992;305:160–4.
- Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–83.
- Kosinski M, Zhao SZ, Dedhiya S, *et al.* Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. *Arthritis Rheum* 2000;43:1478–87.

- 29 Strand V, Singh JA. Newer biological agents in rheumatoid arthritis: impact on health-related quality of life and productivity. *Drugs* 2010;70:121–45.
- 30 EuroQol Group. EQ-5D-5L User Guide. Version 1.0. 2011. http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/UserGuide_EQ-5D-5L.pdf (accessed 1 Jul 2012); http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-5L_UserGuide_2015.pdf (accessed 17 Apr 2017).
- 31 Brooks R. EuroQol: the current state of play. *Health Policy* 1996;37:53–72.
- 32 Herdman M, Gudex C, Lloyd A, *et al*. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;20:1727–36.
- 33 Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics* 1993;4:353–65.
- 34 Smolen JS, Kremer JM, Gaich CL, *et al*. Patient-reported outcomes from a randomised phase III study of baricitinib in patients with rheumatoid arthritis and an inadequate response to biological agents (RA-BEACON). *Ann Rheum Dis* 2017;76:694–700.
- 35 Schiff M, Takeuchi T, Gaich C, *et al*. THU0623 Patient-Reported Outcomes from A Phase 3 Study of Baricitinib in Patients with Early Rheumatoid Arthritis Who Had Received Limited or No Treatment with Disease-Modifying anti-Rheumatic Drugs: Table 1. *Ann Rheum Dis* 2016;75263:419.1–419.
- 36 Emery P, Blanco R, Maldonado Cocco J, *et al*. Patient-reported outcomes from a phase III study of baricitinib in patients with conventional synthetic DMARD-refractory rheumatoid arthritis. *RMD Open* 2017;3:e000410.